D. SIVOUKHINE

COURS DE PHYSIQUE GÉNÉRALE

5 PHYSIQUE ATOMIQUE ET NUCLÉAIRE

PREMIÈRE PARTIE

ÉDITIONS MIR

Д. В. СИВУХИН

ОБЩИЙ КУРС ФИЗИКИ

т. V

АТОМНАЯ И ЯДЕРНАЯ ФИЗИКА

ч. 1

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

D. SIVOUKHINE

COURS DE PHYSIQUE GÉNÉRALE

Tome V

PHYSIQUE ATOMIQUE ET NUCLÉAIRE

Première partie

ÉDITIONS MIR · MOSCOU

Traduit du russe
par SERGUEÏ MEDVÉDEV

На французском языке

AVANT-PROPOS

La compréhension des phénomènes de la physique atomique, et plus généralement de tous les phénomènes du microcosme, doit s'appuyer sur les conceptions de la théorie quantique. C'est pour cette raison que l'auteur a estimé nécessaire d'initier le lecteur aux principaux concepts et principes de la mécanique quantique, bien entendu dans la mesure où on peut le faire dans le cadre d'un cours de physique générale. L'exposé se fonde sur les faits expérimentaux, car c'est la seule voie démontrant qu'il est inéluctable que les conceptions quantiques doivent être adoptées. Cette approche aide d'ailleurs le lecteur à s'habituer au schéma logique et à l'appareil mathématique formel de la mécanique quantique, dont il trouvera un développement systématique dans le cours de physique théorique. On conçoit que dans ces conditions seuls les problèmes simples peuvent être envisagés quantitativement. La majorité des sujets abordés sont étudiés qualitativement en s'appuyant sur les principes généraux.

Dans ce volume, on ne se limite pas seulement à l'étude des phénomènes se produisant dans les enveloppes électroniques des atomes et des molécules. En se conformant à l'ordre de l'évolution historique, on introduit d'abord le concept de photon et on examine les phénomènes optiques liés à ce concept. Ce n'est qu'après qu'on entre dans le vif du sujet de la physique atomique proprement dite — la physique des enveloppes électroniques des atomes. Enfin, on examine brièvement les phénomènes quantiques macroscopiques: la superfluidité, la supraconductibilité, etc.

Comme les tomes précédents, ce livre est une version complétée du cours de physique atomique que l'auteur enseigne depuis de nombreuses années aux étudiants de l'Institut de Physique technique de Moscou.

La deuxième partie de ce tome devait comporter un exposé de la physique nucléaire, mais comme le manuscrit de cette deuxième partie n'était pas encore arrivé à maturation, il a été jugé préférable de consacrer à la physique nucléaire un volume séparé.

Comme dans les tomes précédents, on utilise surtout la forme gaussienne du système d'unités CGS. On utilise aussi l'unité d'énergie hors système qui a fait ses preuves, l'électron-volt (eV) et ses mul-

tiples (keV, MeV, GeV). On n'utilise pas le système international SI. Le caractère artificiel des constantes électrique et magnétique du vide affublées de dimensions rend ce système peu utilisable en physique.

Le membre correspondant de l'Académie des sciences de l'U.R.S.S. S. S. Herstein et le membre correspondant de l'Académie des sciences de l'Ukraine I. S. Gorban, directeur de la chaire de physique expérimentale de l'Université d'Etat de Kiev, ont bien voulu prendre connaissance du manuscrit. L'auteur leur exprime sa vive gratitude pour les remarques qui ont contribué à améliorer ce livre.

D. Sivoukhine

CHAPITRE PREMIER

LES QUANTA DE LUMIÈRE

§ 1. Energie et impulsion du quantum de lumière

1. Vers le milieu du XIXe siècle, on estimait que la nature ondulatoire de la lumière était fermement établie. En témoignaient les phénomènes d'interférences et de diffraction de la lumière. Les expériences de Foucault (1819-1868) excluaient, semblait-il, toute possibilité de retour aux conceptions corpusculaires de la lumière (voir t. IV, § 3). Il en aurait été bien ainsi s'il s'agissait de la théorie corpusculaire newtonienne (voir § 5).

Néanmoins, la théorie ondulatoire de la lumière, même sous sa forme électromagnétique, s'avérait impuissante à expliquer l'ensemble des phénomènes optiques. On s'en rendit compte dès qu'on s'attaqua au problème du *rayonnement d'équilibre* (*rayonnement noir*). Toutes les tentatives de trouver une solution de ce problème, dans le cadre des conceptions ondulatoires sur la base de l'électrodynamique et de la statistique classiques, se soldèrent par un échec (voir t. IV, § 117). Une formule vérifiant les résultats expérimentaux dans toute la gamme des longueurs d'onde fut établie, empiriquement d'abord, par Planck (1858-1947) en octobre 1900. Peu de temps après, Planck trouva une démonstration théorique de sa formule et l'exposa le 14 décembre 1900 à une séance de la Société de Physique d'Allemagne. Ce fait marqua l'apparition de conceptions physiques essentiellement nouvelles — les conceptions *quantiques*. Au début, ces conceptions ne concernaient que la nature de la lumière, mais peu à peu elles s'étendirent à toutes les branches de la physique.

On s'aperçut alors que les notions et les principes de la physique classique, qui avaient été formulés en se fondant sur l'étude d'objets macroscopiques, cessaient d'être valables ou devenaient partiellement valables lorsqu'on descendait à l'échelle atomique ou subatomique. Dans ce domaine d'investigation, de nouvelles conceptions et de nouvelles lois étaient nécessaires et elles furent finalement trouvées. Ces conceptions et ces lois servirent de fondement à une nouvelle physique dite *physique quantique*. Dans un cours de physique générale, on ne saurait exposer la physique quantique d'une façon systématique et déductive, comme on le ferait dans un cours de physique théorique. Il convient, par contre, de retracer ici l'histoire de sa genèse et de son développement, en commençant par plusieurs résultats expérimentaux fondamentaux, dont l'étude permet de dé-

montrer les insuffisances et la validité limitée des conceptions classiques. Ce qui compte surtout, c'est qu'en procédant ainsi on arrive à jalonner la voie qui mène à l'introduction des nouvelles conceptions. C'est cette méthode que nous mettrons en œuvre pour initier le lecteur aux fondements de la physique quantique.

2. Revenons maintenant à la genèse des idées quantiques. Pour établir sa formule du rayonnement d'équilibre, Planck fut amené à introduire une hypothèse complètement étrangère à la physique classique; selon cette hypothèse, *l'émission et l'absorption de la lumière par les substances se produit non pas de façon continue, mais par portions finies ou par quanta.* Comme les propriétés du rayonnement d'équilibre contenu dans une cavité sont indépendantes de la nature de la substance des parois, Planck assimilait la substance à un ensemble d'oscillateurs harmoniques, ce qui n'affectait nullement la généralité des résultats. Afin de concilier cette hypothèse avec les lois de la thermodynamique et de l'électrodynamique, Planck posa que l'énergie $\mathscr{E}$ de chaque quantum rayonné ou absorbé par un oscillateur harmonique de fréquence ν était définie par l'expression suivante:

$$\mathscr{E} = h\nu, \tag{1.1}$$

où h est une constante universelle qui fut appelée ultérieurement *constante de Planck* (voir t. IV, § 118). Planck calcula lui-même la valeur de cette constante à partir des données expérimentales concernant la distribution de l'énergie dans le spectre d'émission du *corps noir.* Mais la constante de Planck, comme toutes les autres constantes fondamentales, intervient dans de nombreux phénomènes physiques, dont l'étude permet de déterminer sa valeur numérique par maints procédés indépendants. Aux erreurs expérimentales près, tous les procédés fournissent des résultats concordants. Selon des données récentes

$$h = 6{,}626176(36)\cdot 10^{-27}\ \text{erg}\cdot\text{s}\ ^{*}). \tag{1.2}$$

Pour certaines applications, notamment en physique théorique, on utilise pour raison de commodité la constante

$$\hbar = 1{,}0545887(57)\cdot 10^{-27}\ \text{erg}\cdot\text{s} \tag{1.3}$$

qui a été introduite par Dirac (1902-1984). On la désigne aussi sous le nom de *constante de Planck* ou de *constante de Dirac.* Avec cette constante, l'énergie du quantum de rayonnement s'exprime par la formule

$$\mathscr{E} = \hbar\omega, \tag{1.4}$$

où $\omega = 2\pi\nu$ est la *fréquence cyclique du rayonnement.* Nous utiliserons aussi bien la formule (1.1) que la formule (1.4).

*) Ici, comme partout ailleurs, on indique entre parenthèses l'écart standard des deux dernières décimales par rapport à leur valeur moyenne.

3. D'après ce qui a été dit plus haut, on voit que Planck pensait que les propriétés quantiques de la lumière ne se manifestaient que *dans les actes d'émission et d'absorption*, autrement dit *lors des interactions de la lumière avec les substances*. Planck estimait, par ailleurs, que la propagation de la lumière dans l'espace était continue et se laissait décrire par les équations classiques de Maxwell (1831-1879). En 1905, Einstein (1879-1955) conféra à la théorie quantique de la lumière une forme plus élaborée. En se fondant sur des considérations théoriques et certaines données expérimentales, Einstein arriva à conclure que même *lors de la propagation dans l'espace la lumière devait se comporter comme un ensemble de particules de nature encore indéterminée*, l'énergie de chaque particule étant définie par la formule de Planck (1.1) ou (1.4). Ultérieurement, ces particules furent appelées *quanta de lumière* ou *photons*.

Ce ne fut pas un simple retour à la théorie corpusculaire de la lumière de Newton. On ne peut en effet considérer les photons comme des particules de lumière ordinaires semblables aux points matériels de la mécanique classique et se déplaçant dans l'espace suivant des trajectoires bien définies. La preuve en est que les photons subissent des *phénomènes d'interférences* et *de diffraction*. Les photons sont doués de propriétés corpusculaires et de *propriétés ondulatoires*. Ce comportement particulier des photons est appelé *dualisme onde-corpuscule*. Il serait dénué de sens de vouloir interpréter le dualisme onde-corpuscule dans le cadre conceptuel classique. L'imagination humaine est impuissante à concevoir une chose qui posséderait simultanément les propriétés d'une onde et celles d'un corpuscule. Mais la Nature est plus riche que l'imagination humaine, et pour l'étudier on doit s'appuyer non sur ce qui est imaginable mais sur les résultats de l'observation et de l'expérience. Notons, en anticipant, que les corpuscules ordinaires, tels les électrons, les neutrons, les atomes, les molécules, etc., possèdent également des propriétés ondulatoires. Nous décrirons au § 18 les expériences qui confirment cette assertion et ce n'est qu'après avoir étudié les propriétés ondulatoires des substances que nous pourrons passer à l'examen de l'interprétation que donne la physique moderne pour le dualisme onde-corpuscule (§ 19).

4. Si le photon possède de l'énergie, il doit avoir une impulsion pour satisfaire à la théorie de la relativité. L'impulsion des photons se manifeste, par exemple, dans l'existence d'une pression lumineuse. Conformément à la théorie de la relativité, l'énergie $\mathscr{E}$ et l'impulsion p d'une particule en mouvement sont liées par la formule

$$\left(\frac{\mathscr{E}}{c}\right)^2 - p^2 = (m_0c)^2 \tag{1.5}$$

(voir t. IV, § 111). (On admet que lors du mouvement de la particule, son état interne et, par suite, sa masse au repos m_0 ne changent pas.) Dans le vide, les photons se déplacent avec la vitesse de la lu-

mière c et sont donc des *particules relativistes*. Si la masse au repos m_0 du photon était différente de zéro, sa masse relativiste serait égale à $m = m_0/\sqrt{1 - v^2/c^2}$, donc infiniment grande. Il s'ensuit qu'on doit poser pour le photon $m_0 = 0$. Par conséquent, la relation ci-dessus doit s'écrire

$$\mathscr{E} = pc. \tag{1.6}$$

(En convenant d'entendre par p l'impulsion du photon dans le sens de propagation de la lumière, on omettra le signe moins dans le calcul de la racine carrée.)

Le caractère non classique de la relation (1.1) ou (1.4) se manifeste, en particulier, en ce que, selon les conceptions classiques, l'énergie doit être liée à l'*amplitude* des vibrations et non pas à leur fréquence. D'après les conceptions classiques, le dualisme onde-corpuscule ne devrait pas exister. Mais si ce dualisme existe, le lien entre les propriétés ondulatoires et corpusculaires du rayonnement ne peut être caractérisé par la seule relation (1.1) ou (1.4). Les propriétés corpusculaires du rayonnement sont caractérisées par l'*énergie* $\mathscr{E}$ et l'*impulsion* $\boldsymbol{p}$, tandis que ses propriétés ondulatoires le sont par la *fréquence* ω et le vecteur d'onde $\boldsymbol{k}$. En théorie de la relativité, $\mathscr{E}$ et $\boldsymbol{p}$ sont réunis en un seul *quadrivecteur énergie-impulsion* $(\mathscr{E}, c\boldsymbol{p})$. Il a été démontré, dans le t. IV, § 107, que la phase $\omega t - \boldsymbol{kr}$ de l'onde est *invariante par rapport à la transformation de Lorentz*. Or, comme $(t, \boldsymbol{r}/c)$ est un quadrivecteur, il en résulte que la fréquence ω et le vecteur d'onde $\boldsymbol{k}$ s'unissent aussi en un *quadrivecteur* $(\omega, c\boldsymbol{k})$. Comme les composantes temporelles $\mathscr{E}$ et ω des quadrivecteurs $(\mathscr{E}, c\boldsymbol{p})$ et $(\omega, c\boldsymbol{k})$ se transforment de la même façon lorsqu'on les soumet aux transformations de Lorentz, la relation (1.1) ou (1.4) satisfait à la condition nécessaire de l'*invariance relativiste*. Or, une relation relativiste ne peut se limiter rien qu'à une relation entre les composantes temporelles des quadrivecteurs $(\mathscr{E}, c\boldsymbol{p})$ et $(\omega, c\boldsymbol{k})$, il doit exister aussi un lien entre les quadrivecteurs eux-mêmes. Il s'ensuit que si l'hypothèse de Planck $\mathscr{E} = \hbar\omega$ est juste, il en découle inévitablement, conjointement avec la condition d'invariance relativiste, que $\boldsymbol{p} = \hbar\boldsymbol{k}$. On adoptera donc par hypothèse que

$$\mathscr{E} = \hbar\omega, \quad \boldsymbol{p} = \hbar\boldsymbol{k} \tag{1.7}$$

Pour un photon cela donne $\mathscr{E}^2 - (pc)^2 = (m_0c^2)^2 = 0$, i.e. la réduction à zéro de la masse au repos m_0 du photon revient à affirmer que pour les photons le quadrivecteur énergie-impulsion $(\mathscr{E}, c\boldsymbol{p})$ est le vecteur lumineux.

5. Par interaction avec les substances, les photons peuvent être émis, absorbés ou diffusés. La conservation du nombre de photons n'a pas lieu, mais les *lois de la conservation de l'énergie et de l'impulsion* doivent être vérifiées.

Un électron libre ne peut que diffuser un photon, sans pouvoir l'émettre ou l'absorber. Pour le démontrer simplement, utilisons un système de référence dans lequel l'électron se trouvait initialement au repos. Posons que l'électron ait émis un photon d'impulsion $\mathbf{p}_{\text{ph}}$ et d'énergie $\mathscr{E}_{\text{ph}}$. Notons $\mathbf{p}_e$ l'impulsion et $\mathscr{E}_e$ l'énergie que possède l'électron après émission du photon. En vertu des lois de la conservation de l'impulsion et de l'énergie on a

$$\mathbf{p}_e + \mathbf{p}_{\text{ph}} = 0, \quad \mathscr{E}_e + \mathscr{E}_{\text{ph}} = m_0c^2,$$

où m_0 est la masse au repos de l'électron. Il s'ensuit

$$(c\mathbf{p}_e)^2 = (c\mathbf{p}_{\text{ph}})^2, \quad \mathscr{E}_e^2 = \mathscr{E}_{\text{ph}}^2 - 2\mathscr{E}_{\text{ph}}m_0c^2 + (m_0c^2)^2.$$

Soustrayons la première égalité de la seconde. Compte tenu de la relation (1.6) pour le photon et de la relation (1.5) pour l'électron, on obtient

$$\mathscr{E}_{\text{ph}}m_0c^2 = 0.$$

Il s'ensuit que $\mathscr{E}_{\text{ph}} = 0$, ce qui signifie qu'il ne peut y avoir d'émission. Le même raisonnement permet de démontrer que l'absorption d'un photon est tout aussi impossible.

Ce résultat est trivial, puisque la démonstration impliquait tacitement que la masse au repos de l'électron avant émission d'un photon était égale à sa masse au repos après émission. Or, cela signifie que *l'état interne de l'électron ne s'est pas modifié du fait de l'émission.* Dans ces conditions, l'énergie totale de l'électron ne pourrait qu'augmenter grâce à l'énergie cinétique qu'il acquiert par recul lors de l'émission du photon. Le photon émis devrait emporter une énergie positive. En conséquence, si l'électron pouvait émettre un photon, la loi de la conservation de l'énergie aurait été violée.

Pour conclure ce paragraphe, précisons les notations utilisées. En physique des particules élémentaires, on entend par masse de toute particule sa *masse au repos* que l'on dénote m (en omettant l'indice 0). Néanmoins, comme on aura à utiliser, en plus de la masse au repos, la masse relativiste des particules, nous conserverons pour les dénoter les notations usuelles, i.e. m_0 et m. Ces notations seront utilisées plus particulièrement chaque fois que l'espèce de la particule n'aura pas à être précisée. Lorsqu'il s'agira de particules définies, telles que électrons, protons, neutrons, etc., m_e, m_p, m_n, ... dénoterons leurs masses au repos. Dans les considérations de la partie générale, pour raison de commodité, on dénotera la masse au repos par m_0 afin de la distinguer de la masse relativiste m. Dans les résultats définitifs, il sera utile de remplacer m_0 par m_e, m_p, m_n, . . ., afin de bien mettre en évidence à quelle particule se rapporte la masse au repos.

PROBLÈMES

1. Calculer la longueur d'onde λ_C, pour laquelle l'énergie d'un quantum de lumière est égale à l'énergie au repos de l'électron. Cette longueur d'onde est appelée *longueur de Compton* de l'électron (voir § 3).

Réponse. $\lambda_C = \frac{h}{m_0 c} \approx 2{,}43 \cdot 10^{-10}$ cm, où m_0 est la masse au repos de l'électron.

2. Calculer la masse relativiste m d'un quantum de lumière de longueur d'onde λ.

Réponse. $m = \frac{\lambda_C}{\lambda} m_0$.

3. Si le photon possédait une masse au repos égale à m_{ph}, la vitesse de la lumière dans le vide devrait dépendre de la longueur d'onde. Par étude expérimentale de cette dépendance on pourrait évaluer la limite inférieure de la masse du photon. Trouver les expressions de la vitesse de phase et de la vitesse de groupe de la lumière dans le vide en admettant que $m_{ph} \neq 0$.

Solution. L'énergie $\hbar\omega$ et l'impulsion $\hbar k$ du photon sont liées par la relation

$$(\hbar\omega)^2 - (c\hbar k)^2 = (m_{ph} c^2)^2, \tag{1.8}$$

où c désigne non plus la vitesse de la lumière dans le vide, mais une certaine vitesse fondamentale utilisée dans la théorie de la relativité. L'équation (1.8) permet de trouver la vitesse de phase $v = \omega/k$ de la lumière dans le vide:

$$\frac{1}{v} = \frac{1}{c}\sqrt{1 - \left(\frac{m_{ph}c^2}{\hbar\omega}\right)^2} \approx \frac{1}{c}\left(1 - \frac{1}{2}\left(\frac{m_{ph}c^2}{\hbar\omega}\right)^2\right). \tag{1.9}$$

Dans les expériences, on mesure non pas la vitesse de phase, mais la vitesse de groupe de la lumière. On tire de (1.8)

$$u = \frac{d\omega}{dk} = c^2\frac{k}{\omega} = \frac{c^2}{v}, \tag{1.10}$$

ou

$$u = c\left[1 - \frac{1}{2}\left(\frac{m_{ph}c^2}{\hbar\omega}\right)^2\right] = c\left[1 - \frac{1}{2}\left(\frac{c\lambda}{h} m_{ph}\right)^2\right]. \tag{1.11}$$

Les méthodes de radiodétection modernes servant à mesurer la vitesse de la lumière à différentes fréquences fournissent le résultat suivant: $m_{ph} < 4 \times 10^{-21} m_e$, m_e étant la masse de l'électron.

§ 2. Effet photoélectrique

1. L'un des phénomènes confirmant l'hypothèse des photons est l'*effet photoélectrique* que nous étudierons maintenant.

En 1887, Heinrich Hertz (1857-1894) constata que lorsqu'on éclairait avec de la lumière ultraviolette l'*électrode négative* d'un éclateur à étincelle sous tension, le passage des étincelles s'en trouvait facilité. Hertz, absorbé par l'étude des ondes électromagnétiques prévues par Maxwell, n'accorda pas d'attention à ce phénomène. Les premières études de ce phénomène furent entreprises par Hallwachs (1859-1922), Righi (1850-1920) et surtout par A.G. Stoletov (1839-1896).

Le phénomène que découvrit Hertz consiste en ce que *tout corps métallique portant des charges négatives perd celles-ci lorsqu'on l'éclaire avec de la lumière ultraviolette.* Si on irradie avec le même rayonnement un corps chargé positivement, la perte de charge n'est pas observée. D'autre part, si le corps métallique est initialement élec-

triquement neutre, il se charge *positivement* jusqu'à un potentiel de plusieurs volts dès qu'on l'éclaire. Après la découverte, en 1897, de l'électron par J.J. Thomson (1856-1940), ce dernier ainsi que Lenard (1862-1947) réussirent à mesurer la charge massique *e/m* des particules qui s'échappent des corps lorsqu'on les éclaire. La valeur de ce rapport fut trouvée égale à celle qui fut déterminée pour les particules constituant les rayons cathodiques. On démontra ainsi que les corps irradiés par un rayonnement lumineux perdent des *électrons.*

Le phénomène consistant en une expulsion d'électrons par les substances soumises à l'action de la lumière fut dénommé *effet photoélectrique.* On distingue l'effet photoélectrique *extérieur* et l'effet photoélectrique *intérieur.* Dans l'effet photoélectrique extérieur, sous l'action de la lumière incidente ce sont les électrons de la couche périphérique du corps qui sont arrachés et s'échappent dans le milieu adjacent, le vide par exemple. Dans le cas de l'effet photoélectrique intérieur, les électrons excités par la lumière restent à l'intérieur du corps éclairé dont la neutralité électrique n'est pas compromise. Pour la justification de l'hypothèse photonique, c'est l'effet photoélectrique extérieur qui importe surtout et c'est cet effet que nous étudierons en premier lieu ; à la fin de ce paragraphe, on donnera quelques indications sur l'effet photoélectrique intérieur et les applications.

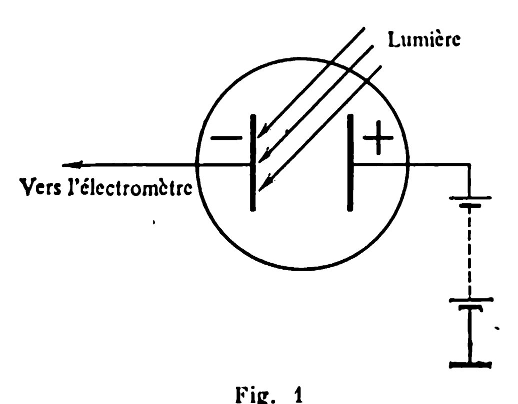


Fig. 1

Les électrons qui sont arrachés par la lumière incidente sont appelés *photoélectrons.* L'effet photoélectrique se manifeste dans les métaux, les diélectriques, les semiconducteurs et les électrolytes ; pour que l'effet puisse se manifester, la condition nécessaire (mais non suffisante) est que la lumière utilisée subisse une absorption suffisante dans la couche superficielle du corps éclairé. L'effet photoélectrique est provoqué non seulement par les rayons ultraviolets. Les métaux alcalins tels que lithium, sodium, potassium, rubidium et césium, sont également fort sensibles à l'action photoélectrique dans la zone visible du spectre. Si on soumet la surface de ces métaux, ainsi que de certains autres, à un traitement spécial, l'effet photoélectrique pourra se manifester même sous l'action des rayons infrarouges.

2. La figure 1 représente le schéma de principe du montage utilisé pour l'étude expérimentale de l'effet photoélectrique. Les pho-

toélectrons arrachés de la cathode par suite de son éclairement par une lumière convenable sont entraînés par la tension appliquée vers l'anode et ferment le circuit. En mesurant la vitesse de charge de l'électromètre (que l'on peut remplacer par un galvanomètre sensible), on détermine l'intensité du courant électrique passant dans le circuit et on en déduit le nombre de photoélectrons parvenant à l'anode dans l'unité de temps. Dans les premières études expérimentales, les expériences de ce type étaient réalisées dans les gaz, mais il est préférable de les réaliser dans le vide, car les gaz compliquent les phénomènes évoluant dans la couche superficielle du métal.

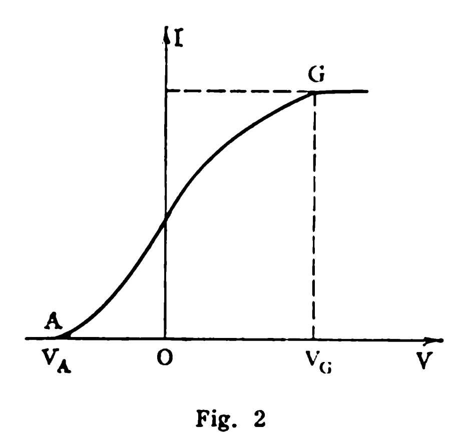


Fig. 2

Pour une substance donnée, l'effet photoélectrique dépend fortement de l'état de sa surface. Pour que les résultats soient univoques, l'étude de l'effet doit être réalisée dans un vide poussé et la surface du corps étudié doit être préalablement parfaitement décapée. Le matériau servant de cathode doit être purifié par distillation dans le vide, puis déposé sous forme d'une couche mince sur un support quelconque.

En maintenant constantes l'intensité et la fréquence de la lumière incidente, faisons varier la tension V appliquée entre l'anode et la cathode. La dépendance du photocourant I avec la tension V sera alors traduite par une courbe représentée sur la figure 2. Cette courbe est la *caractéristique de la cellule photoélectrique*, i.e. du dispositif permettant de produire l'effet photoélectrique. Lorsqu'on fait croître la tension appliquée, la caractéristique se transforme en une droite horizontale, correspondant au courant maximal, que l'on appelle *courant de saturation*. On atteint à la saturation lorsque tous les électrons arrachés par la lumière incidente de la surface de la cathode parviennent à l'anode. Une augmentation ultérieure de la tension ne peut faire varier l'intensité du photocourant puisque *celle-ci ne dépend que du nombre d'électrons arrachés de la cathode par seconde.*

L'existence du courant de saturation a été expérimentalement établie par A. G. Stoletov. Ce fut encore lui qui démontra que *l'intensité du courant de saturation était rigoureusement proportionnelle à l'intensité de la lumière incidente*, à condition que sa fréquence reste constante. Cela signifie que le courant de saturation est proportionnel à la puissance de la *lumière absorbée.* Mais comme cette dernière est proportionnelle à son tour à l'intensité de la lumière incidente,

c'est cette dernière qui figure dans l'énoncé de la loi de proportionnalité. La loi de Stoletov ne s'applique rigoureusement que si le courant de saturation est constitué uniquement par les électrons libérés par la lumière sur la *surface photosensible du corps* étudié. A cet effet il est indispensable de placer la surface photosensible dans le vide. Dans les dispositifs remplis de gaz, on peut observer des écarts sensibles par rapport à la simple proportionnalité. Les dispositifs à gaz ont généralement une sensibilité plus grande que les appareils à vide, car au courant d'électrons arrachés par la lumière sur la surface photosensible du corps vient s'ajouter le *courant d'ionisation du gaz* contenu dans le dispositif. Lorsqu'on utilise des dispositifs à remplissage gazeux pour effectuer des mesures, on doit tenir compte de cette dernière circonstance.

3. On pourrait essayer de trouver une explication qualitative de l'effet photoélectrique sur la base de la théorie ondulatoire de la lumière. Les électrons libres, qui existent en grande quantité dans les métaux, y sont retenus par le champ retardateur qui existe à la frontière du métal. Le travail de sortie d'un électron du métal est de l'ordre de quelques électrons-volts. Soumis à l'action du champ électrique de l'onde lumineuse, l'électron libre se met à osciller. Lorsque l'énergie de ces oscillations devient suffisante, l'électron peut surmonter le champ qui s'oppose à son extraction et s'échapper du métal. S'il s'agit d'électrons liés, la situation sera semblable, à cela près que la dépendance de l'énergie d'oscillations avec la fréquence sera plus compliquée et présentera un caractère de résonance.

Or, cette explication ne cadre pas avec les données expérimentales. L'énergie cinétique de l'électron oscillant doit être empruntée à l'onde lumineuse. Il s'ensuivrait que l'énergie de l'électron éjecté devrait être d'autant plus grande que l'intensité de la lumière incidente est grande. Or, les expériences de Lenard et de nombreux autres scientifiques ont montré que *la vitesse maximale des électrons sortant d'un corps est indépendante de l'intensité de la lumière incidente et ne dépend que de sa fréquence.* (On admet que la substance du corps et l'état de sa surface éclairée restent inchangés.) Cette règle est violée pour l'effet photoélectrique non linéaire (cf. pt. 9).

L'explication ci-dessus conduit à un désaccord flagrant avec l'expérience lorsqu'on cherche à évaluer le *temps d'établissement de l'effet photoélectrique.* Prenons, par exemple, une lampe d'éclairage électrique d'une puissance $P = 100$ W. Posons pour simplifier que c'est une source lumineuse ponctuelle, émettant la lumière de façon isotrope et continue dans toutes les directions. Selon la théorie ondulatoire classique, le flux d'énergie radiante se propage *continuellement* de la source dans toutes les directions. Posons que la lumière tombe sous incidence normale sur une photocathode en zinc plane, se trouvant à une distance r de la source lumineuse. La luminence énergétique créée par la lampe sur la photocathode est alors égale à

$P/(4\pi r^2)$. Pour le zinc le travail A d'extraction des électrons est d'environ 3,74 eV. Il tombe sous le sens qu'un électron ne pourra s'échapper du métal que si l'énergie qu'il acquiert lors de ses oscillations forcées est égale ou supérieure à A. L'énergie maximale qu'un atome peut soustraire au rayonnement pendant le temps t est $\mathcal{E}_{max} = \frac{P}{4\pi r^2}\sigma t$, où σ est l'aire de la section droite de l'atome. L'énergie transférée à l'électron pendant ce même temps est inférieure à $\mathcal{E}_{max}$, mais elle ne doit cependant pas être inférieure à A si on veut que l'électron s'échappe du métal, ce qui implique que $\mathcal{E}_{max} > A$. On en déduit que $t > \frac{4\pi r^2}{\sigma P} A$. La distance moyenne entre les atomes de zinc est donnée par la formule $\left(\frac{M}{\delta A_{Av}}\right)^{1/3}$, où A_{Av} est le nombre d'Avogadro, M la masse atomique relative et δ la densité du zinc. Pour le zinc $M = 65$, $\delta = 7$ g/cm³, et, par conséquent, $d = 2{,}49 \cdot 10^{-8}$ cm. On évalue la section σ par la formule $\sigma = d^2 \approx 6 \cdot 10^{-16}$ cm². En posant $r = 1$ m, on trouve $t \geqslant 1{,}25$ s. Ce résultat montre que *d'après la théorie ondulatoire classique l'effet photoélectrique doit se manifester avec un certain retard.* En remarquant que les cellules photoélectriques réagissent à des flux lumineux notablement plus faibles que celui qu'on a envisagé plus haut, le retard devrait être beaucoup plus grand que celui que nous venons de calculer. Or, *l'expérience montre que l'effet photoélectrique ne présente aucune inertie, i.e. le courant s'établit instantanément, simultanément avec l'éclairement de la cellule.* C'est justement cette réponse instantanée des cellules photoélectriques qui détermine la majorité de leurs applications dans les sciences et les techniques.

4. Toutes les difficultés mentionnées disparaissent dès qu'on fait appel à la théorie photonique. Lorsqu'un photon entre en interaction avec un électron du métal, il se produit entre eux un échange d'énergie et d'impulsion. Comme le processus d'interaction ressemble aux chocs de billes, on l'appelle processus de *collision.* L'effet photoélectrique résulte de *collisions non élastiques* entre les photons et les électrons. Lors de ces collisions les photons sont absorbés par les électrons, qui captent ainsi l'énergie des photons incidents. Ainsi l'énergie de l'électron augmente non pas progressivement mais *d'un seul coup, au cours d'une seule collision*, et c'est pour cela que l'effet photoélectrique ne présente pas d'inertie.

L'énergie du photon absorbé peut être utilisée pour arracher un électron d'un atome appartenant au métal. L'électron arraché peut interagir avec un des atomes du métal et dissiper son énergie en chaleur. Un électron s'échappant du métal possède le maximum d'énergie s'il était *libre* à l'intérieur du métal, i.e. n'était pas lié à un atome et s'il ne dissipe pas de l'énergie en chaleur en quittant le métal. Dans ces conditions, l'électron ne dépense son énergie cinétique que

pour surmonter les forces qui s'opposent à sa sortie à la périphérie du métal, i.e. pour réaliser le travail d'extraction. Supposons que l'électron ait acquis de l'énergie cinétique par collision *avec un seul photon*. Les processus à plusieurs photons, que nous étudierons au sous-paragraphe 9, quoique réalisables, sont peu probables aux faibles intensités lumineuses (i.e. en optique linéaire). Dans ce cas, *l'énergie cinétique maximale que peut avoir un électron éjecté du métal est donnée par la formule*

$$1/2\, m_e\, v_{max}^2 = h\nu - A, \tag{2.1}$$

où A est le travail d'extraction et m_e la masse au repos de l'électron. Cette formule fut établie par Einstein et porte son nom.

Avant de passer à la discussion de la formule d'Einstein, il importe de préciser de quelle manière « un électron libre contenu dans le métal » peut absorber un photon, et de se rendre compte que cette éventualité ne contredit pas l'assertion démontrée à la fin du paragraphe précédent, selon laquelle l'absorption d'un photon par un électron libre est incompatible avec les lois de la conservation de l'énergie et de l'impulsion. En fait, il n'y a là aucune contradiction, parce que cette dernière n'apparaît que par suite d'une terminologie inadéquate. En réalité, « un électron libre contenu dans le métal » *n'est pas libre* au sens propre du terme. Cet électron est comme enfermé dans une boîte sur les parois de laquelle agit un champ retardateur. Le photon interagit, d'une part, avec l'électron et, d'autre part, l'électron et le photon interagissent avec le métal tout entier. Or, lorsque trois corps entrent en interaction, les lois de la conservation de l'énergie et de l'impulsion peuvent être satisfaites simultanément. L'impulsion du photon est transmise simultanément à l'électron et au métal tout entier, tandis que l'énergie n'est transmise qu'à l'électron, vu que la masse du métal peut être posée infiniment grande.

5. De la formule d'Einstein (2.1) découlent deux corollaires rigoureusement confirmés par l'expérience.

1. *L'énergie cinétique maximale des électrons extraits du métal dépend linéairement de la fréquence de la lumière d'éclairage et ne dépend pas de son intensité. L'intensité de la lumière influe sur la quantité d'électrons éjectés, mais n'exerce aucune action sur leur énergie cinétique maximale.* Il est intéressant de noter que la tangente de l'angle que fait la droite (2.1) avec l'axe des fréquences ν est égale à la constante de Planck h (on porte en abscisses la fréquence ν et en ordonnées l'énergie cinétique maximale de l'électron $1/2\ m_e v_{max}^2$). On dispose ainsi d'une nouvelle méthode de mesure de la constante de Planck.

2. *Il existe une fréquence limite inférieure (dite fréquence de seuil) ν_0 donnant naissance à l'effet photoélectrique. Au-dessous de cette fréquence limite, l'effet ne se manifeste pas.* La valeur de cette fréquence

limite dépend de la composition du corps irradié et de l'état de sa surface.

En mettant le travail d'extraction sous la forme $A = h\nu_0$, où ν_0 est une constante positive, écrivons la formule (2.1) de la façon suivante :

$$\frac{1}{2} m_e v_{max}^2 = h(\nu - \nu_0) \quad (2.2)$$

Pour $\nu < \nu_0$ le second membre est négatif ; comme le premier membre est une quantité essentiellement positive, ce cas doit être exclu, ce qui implique l'absence de l'effet photoélectrique. La fréquence ν_0 est la *limite basse fréquence de l'effet photoélectrique* (*seuil photoélectrique*). Du point de vue ondulatoire, on n'arrive pas à expliquer l'existence de ce seuil.

6. La vérification expérimentale de la formule d'Einstein et de ses corrollaires joue un rôle de premier plan non seulement pour la théorie de la lumière, mais pour la physique toute entière. Elle consiste à relever les caractéristiques des cellules photoélectriques où l'éclairement excite un courant électrique. Le schéma du montage correspondant est représenté sur la figure 1. Il importe de pouvoir mesurer le photocourant non seulement lorsque la différence de potentiel, appliquée entre la cathode et l'anode, est positive (champ accélérateur), mais aussi lorsque cette différence de potentiel est *négative* (champ ralentisseur). Pour rehausser la précision des mesures, le dispositif doit être conçu de façon que tous les photoélectrons arrachés de la cathode par la lumière parviennent à l'anode. Pour y arriver on remplace le condensateur plan par un *condensateur sphérique*, dont l'armature externe est constituée par une sphère métallique de grand diamètre, et l'armature interne est une petite bille fabriquée avec la substance à étudier.

Tout photoélectron libéré par la lumière peut entrer en collision avec un des atomes de la couche périphérique du métal. De ce fait il peut être ralenti ou même ne pas s'échapper du métal. Le champ électrique appliqué accélère les électrons freinés et facilite leur sortie du métal. C'est pour cela que l'intensité du photocourant augmente avec la tension appliquée entre la cathode et l'anode. Pour vérifier l'exactitude de la formule d'Einstein (2.1) on doit mesurer l'énergie cinétique maximale que peut avoir un photoélectron à la sortie du métal. Une telle mesure est rendue difficile par l'existence d'une *différence de potentiel de contact* qui est très sensible au traitement superficiel du métal et qui, de ce fait, est difficile à contrôler. Afin d'éliminer l'influence qu'exerce la différence de potentiel de contact, on repère sur la caractéristique tension-courant des cellules photoélectriques deux points : le point A où le courant est annulé par l'application d'un champ retardateur et le point G où apparaît le courant de saturation (fig. 2). Notons V_A et V_G les indications du voltmètre correspondant à ces points. Il s'agit bien d'*indications du voltmètre*

et non pas de différences de potentiel entre anode et cathode, puisqu'on doit ajouter à celles-ci la différence de potentiel de contact V_c entre ces mêmes électrodes. Les différences de potentiel totales sont respectivement égales à $V_A + V_c$ et $V_G + V_c$.

Lorsque le voltmètre indique une tension inférieure à V_A, tous les électrons sont bloqués et aucun ne parvient à l'anode. Lorsque le voltmètre indique la valeur V_A, ne parviennent à l'anode que les électrons possédant la vitesse maximale v_{max}. En égalant les énergies que possèdent les électrons aux points A et G de la caractéristique, on obtien

$$^1/_2\, m_e v_{max}^2 = -e\,(V_A + V_c),$$

où e désigne la valeur absolue de la charge de l'électron. (On néglige ici les vitesses thermiques, vu que l'énergie moyenne kT du mouvement thermique est négligeable devant l'énergie $h\nu$ des photons.) Lorsque le voltmètre indique la valeur V_G, même les électrons sortant avec une vitesse nulle parviennent à l'anode, et par suite

$$0 = -e\,(V_G + V_c).$$

On en tire la différence de potentiel de contact: $V_c = -V_G$. La première équation permet de trouver l'énergie cinétique maximale des photoélectrons:

$$^1/_2\, m_e v_{max}^2 = e\,(V_G - V_A). \qquad (2.3)$$

Il résulte de l'équation $V_G = -V_c$ que la position du point V_G sur la caractéristique tension-courant de la cellule considérée ne dépend que de la différence de potentiel de contact et ne dépend donc nullement de la fréquence ν de la lumière incidente. Par contre, la position de point V_A change avec la fréquence ν, puisque l'énergie maximale $1/2\ m_e v_{max}^2$, dont dépend la quantité $V_A + V_c$, dépend de la fréquence. Lorsque la fréquence de la lumière augmente, les caractéristiques se déplacent vers la gauche, mais l'abscisse du point V_G reste fixe. Si on trace les caractéristiques à une échelle telle que le courant de saturation soit toujours représenté par des ordonnées de même valeur, toutes les caractéristiques passeront par un seul et même point G.

Au début du sous-paragraphe 6, on avait mentionné que la forme sphérique du condensateur était préférable à la forme plane. Maintenant que nous avons appris à connaître le principe de la méthode de mesure de l'énergie cinétique maximale des photoélectrons, nous pouvons indiquer un autre avantage de la forme sphérique. Dans le cas d'un condensateur plan, pour que les photoélectrons parviennent à l'anode, ce ne sont pas leurs vitesses totales, mais les composantes normales à la surface de l'anode qui comptent. L'existence de composantes tangentielles fait qu'aux points A et G la caractéristique de la cellule, au lieu de couper les droites horizontales, s'en approche de

manière *asymptotique*. Il devient alors difficile de déterminer les positions des points A et G. La situation est tout autre dans le cas d'un condensateur sphérique à bille intérieure de petit diamètre, puisque les vitesses des photoélectrons sont pratiquement toutes radiales, i.e. normales à la surface de l'anode.

Une vérification expérimentale précise de la formule d'Einstein fut entreprise d'abord en 1912 par Richardson (1879-1959) et Carl Compton (1887-1954), puis en 1916 des expériences encore plus pré-

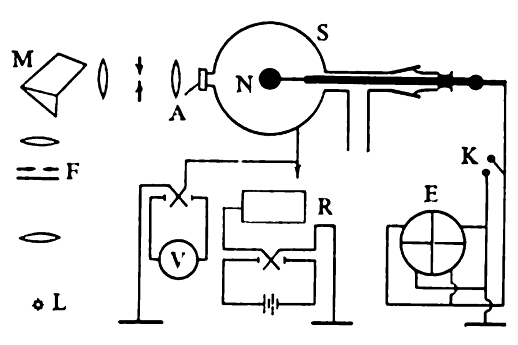


Fig. 3

cises furent réalisées par Millikan (1868-1953). Ces deux cycles de recherches confirmèrent l'exactitude de la formule d'Einstein, quoiqu'elles ne fussent pas irréprochables. Une nouvelle vérification de cette formule, réalisée dans des conditions de très grande propreté expérimentale, fut entreprise en 1926 par P. I. Loukirski (1894-1954) et S. S. Prilejaev. Le schéma du montage utilisé est représenté sur la figure 3. Le dispositif se présente sous la forme d'un condensateur sphérique, à l'intérieur duquel on fait le vide; l'armature extérieure est constituée par une couche d'argent déposée sur la surface intérieure de la sphère S (d'un diamètre de 11 cm). L'armature intérieure du condensateur était constituée par une bille N (de 1,5 cm de diamètre) fabriquée avec le métal étudié; c'est la surface de cette bille qui servait à l'étude de l'effet photoélectrique. Pour rendre aisé le remplacement des billes, on les vissait sur une tige solidaire de l'un des verres rodés. Pour éliminer toute contamination par les charges électriques provenant des parties en verre, on utilisait une isolation en quartz. La bille N était éclairée à travers une fenêtre A en verre de quartz. La lumière produite par un arc à mercure était dirigée sur le monochromateur M. On détermina l'effet photoélectrique généré par les métaux: Al, Zn, Sn, Ni, Ag, Cd, Pb, Cu, Pt. La figure 4 représente les caractéristiques de la cellule pour le zinc ($\lambda = 230{,}2$; 253,7 ; 313,0 nm respectivement pour les courbes *1*, *2*, *3*). Pour tous les autres métaux étudiés les caractéristiques avaient une forme

analogue. Toutes les caractéristiques étaient tracées à une échelle où le courant de saturation était posé conventionnellement égal à 100. Sous cette condition, pour un métal donné, toutes les courbes relevées à des longueurs d'onde différentes se coupaient en haut en un même point, au-delà duquel elles se transformaient en droites horizontales. Ce résultat est conforme aux considérations développées

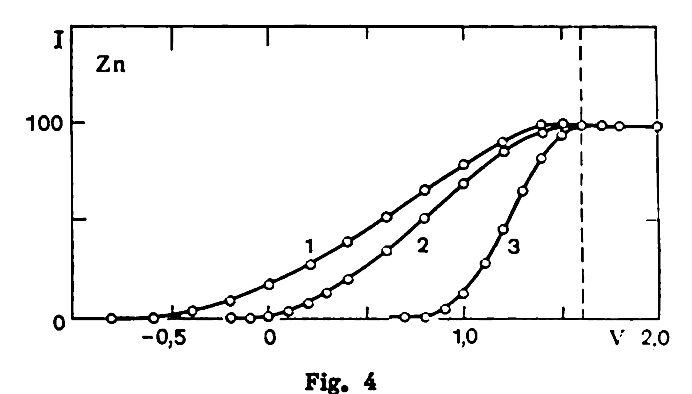


Fig. 4

plus haut. Les auteurs de ces expériences ont trouvé pour la constante de Planck la valeur $h = 6{,}58 \cdot 10^{-27}$ erg·s (on a recalculé leur résultat pour tenir compte de la valeur moderne de la charge de l'électron). (Attention! Sur la figure 3 l'extrémité droite du potentiomètre doit être mise à la terre.)

7. Dans le cas des rayons X, l'énergie du quantum de lumière $h\nu$ est grande par rapport au travail de sortie A. Cela revient à dire que la fréquence ν est grande devant le seuil photoélectrique ν_0. Dans ce cas, dans (2.1), on peut négliger A et cette équation s'écrit alors

$$^1/_2\, m_e v^2_{max} = h\nu, \tag{2.4}$$

ou

$$eV = h\nu, \tag{2.5}$$

où eV est l'énergie de l'électron exprimée en fonction de la tension d'accélération. Cette formule s'applique également à l'*effet photoélectrique normal*, où l'énergie des quanta de lumière est transformée en énergie cinétique des électrons et à l'*effet photoélectrique inverse* dans lequel des quanta de rayons X sont produits aux dépens de l'énergie cinétique des électrons bombardant le métal. C'est ce processus qui est mis en œuvre dans les tubes à rayons X. Dans ce dernier cas, la formule (2.5) détermine la *fréquence maximale* des rayons X émis par l'anticathode pour une tension donnée appliquée au tube. L'existence d'une fréquence limite maximale est confirmée par l'expérience. Ce fait constitue une nouvelle preuve de la justesse de la théorie quantique de la lumière.

Lorsque les électrons tombent sur l'anticathode, ils subissent un freinage brusque, ce qui fournit un *rayonnement* X *de freinage*. Le spectre de ce rayonnement, tracé en fonction des longueurs d'onde, est *continu*, tout comme celui de la lumière blanche. Aussi l'appelle-t-on *rayonnement* X *blanc*. Son intensité spectrale I_λ pour différentes valeurs de la tension appliquée au tube est représentée sur la figure 5 (pour une anticathode en tungstène). Du côté des grandes longueurs d'onde, la courbe d'intensité décroît lentement et tend asymptotiquement vers zéro à mesure que la longueur d'onde augmente. Du côté des petites longueurs d'onde, la courbe d'intensité s'arrête brusquement pour une certaine longueur d'onde appelée *limite inférieure des longueurs d'onde du spectre continu des rayons* X. Cette limite est donnée par la formule (2.5) d'où il découle

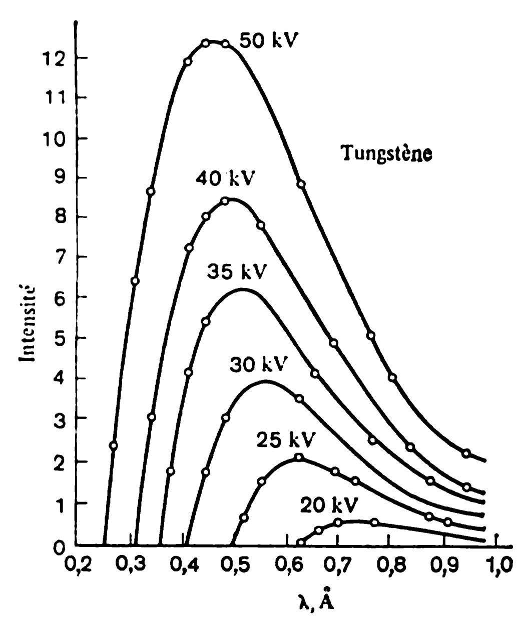


Fig. 5

$$\lambda = \frac{hc}{eV} = \frac{1,240}{V}\,\text{nm}, \quad (2.6)$$

la tension V appliquée au tube étant mesurée en kilovolts. Cette limite inférieure ne dépend pas de la nature de l'anticathode et dépend uniquement de la *tension appliquée au tube*. Si la tension appliquée dépasse une certaine limite dépendant de la nature de l'anticathode, sur le spectre continu viennent se superposer des raies spectrales fines dont l'ensemble constitue le *spectre de raies caractéristiques*, propre au métal constituant l'anticathode. Mais même dans ce cas, la limite du côté des ondes courtes subsiste et est toujours définie par la formule (2.6). L'existence de cette limite constitue une des manifestations les plus marquantes des propriétés corpusculaires du rayonnement X.

La mesure de la limite inférieure du rayonnement X est une méthode précise de détermination de la constante de Planck h, à l'aide de la formule (2.6). On utilise cette même formule dans la mesure des longueurs d'onde du rayonnement X dur et du rayonnement gamma.

8. Jusqu'à présent, dans notre description de l'effet photoélectrique, nous avons surtout cherché à mettre en relief les propriétés *corpusculaires* de la lumière. Cependant les photons sont doués également de propriétés ondulatoires. On peut se demander si ces propriétés ondulatoires se manifestent dans l'effet photoélectrique. Les propriétés ondulatoires des photons se manifestent dans les lois régissant l'effet photoélectrique dit *sélectif*. Notons I_λ l'intensité du courant de saturation rapportée à un intervalle unité de longueurs d'onde

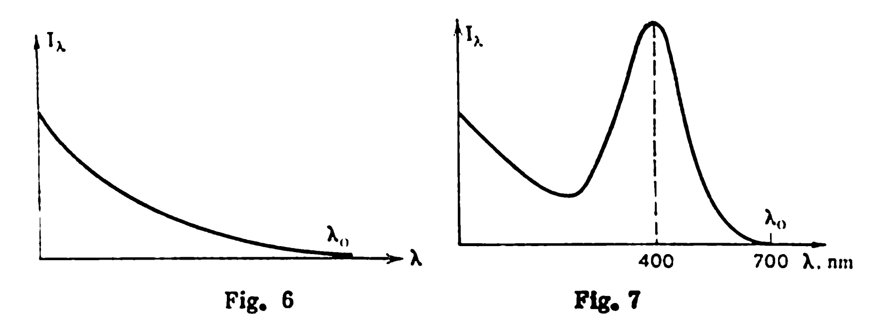


Fig. 6 Fig. 7

et à l'unité d'énergie radiante absorbée. Dans le cas de l'effet photoélectrique *normal*, la valeur de I_λ décroît de façon monotone avec la longueur d'onde et tend vers zéro à la limite située du côté des grandes longueurs d'onde $\lambda = \lambda_0$ (fig. 6). Lorsque le vecteur électrique de l'onde incidente *n'est pas normal* au plan d'incidence, on observe que la courbe $I_\lambda(\lambda)$ pour différents métaux, plus particulièrement pour les métaux alcalins, présente un *maximum aigu* dans une partie déterminée du spectre. Par exemple, pour le sodium le maximum se situe à $\lambda = 340$ nm, pour le potassium à $\lambda = 435$ nm, pour l'aluminium à $\lambda = 215$ nm, etc. (voir la courbe de la figure 7 qui se rapporte à un alliage de potassium et de sodium). Il est vraisemblable que les autres métaux présentent des maximums analogues, qui se situent loin dans la gamme des petites longueurs d'onde et sont donc difficilement observables. L'effet photoélectrique, dont la courbe $I_\lambda(\lambda)$ présente des maximums de ce type, est appelé *effet photoélectrique sélectif*.

Le maximum de la courbe $I_\lambda(\lambda)$ ressemble au maximum correspondant à la résonance des oscillations forcées d'un oscillateur harmonique. On peut en déduire que dans l'effet photoélectrique sélectif les électrons des métaux se comportent comme s'ils possédaient des *fréquences propres*, à proximité desquelles s'observent les maximums de la grandeur $I_\lambda(\lambda)$.

Une autre particularité de l'effet sélectif où se manifestent aussi les propriétés ondulatoires de la lumière réside en ce que *l'intensité de la photoémission dépend fortement de la polarisation de la lumière*

incidente ainsi que de l'angle d'incidence. L'effet sélectif est absent lorsque le vecteur électrique de l'onde incidente est perpendiculaire au plan d'incidence, ou bien si la lumière tombe normalement. Dans les deux cas, la lumière incidente comporte une composante du champ électrique qui est tangente à la frontière du métal, mais ne comporte pas de composante normale. Le phénomène se produit comme si la sélectivité de l'effet était déterminée par la composante du vecteur électrique *normale à la surface du métal.* La composante normale arrache les électrons du métal plus efficacement que ne le fait la composante tangentielle. Le caractère sélectif de l'effet est particulièrement notable pour une lumière polarisée rasante dont le vecteur électrique est contenu dans le plan d'incidence. Ces assertions sont illustrées par la figure 8 représentant la dépendance de I_λ avec λ pour différentes polarisations de la lumière incidente, et par la figure 9 illustrant l'influence de l'angle d'incidence. Ces résultats ont été obtenus avec le même alliage sodium-potassium qui a servi à relever la courbe de la figure 7.

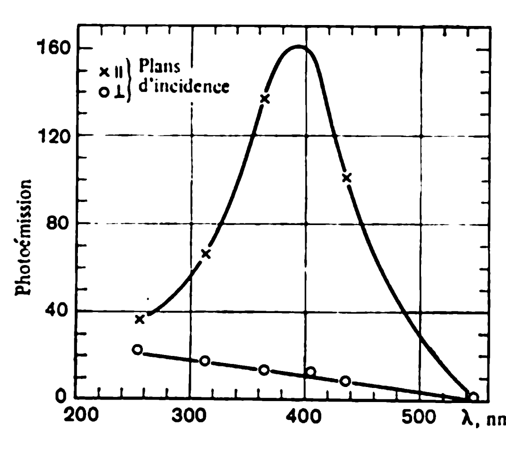


Fig. 8

9. Depuis l'apparition des lasers il est devenu possible de procéder à des expériences avec des faisceaux lumineux de grande puissance et d'observer un comportement *non linéaire* de l'effet photoélectrique. Lorsque l'intensité de la lumière devient suffisamment grande, l'électron, avant de quitter la cathode, peut entrer en collision *avec plusieurs photons* (*processus à plusieurs photons*). Dans ce cas, l'équation (2.1) doit être remplacée par l'équation

$$^1/_2\, m_e v_{max}^2 = Nh\nu - A, \tag{2.7}$$

où N est le nombre de collisions de l'électron avec les photons. Le seuil photoélectrique existe toujours, mais la fréquence de seuil est donnée maintenant par l'expression $\nu_0 = A/(hN)$, ce qui signifie qu'elle est N fois plus petite que dans l'effet à photon unique. Pendant longtemps l'échauffement de la cathode, produit par le faisceau laser, s'opposait à l'observation de l'effet non linéaire, car il s'accompagnait d'*émission thermo-ionique*, qui ne présente, naturellement, aucun seuil basse fréquence. On arriva à éliminer presque complètement l'influence de l'émission thermo-ionique en utilisant

des *impulsions laser ultra-brèves* — d'une durée de 10^{-11} à 10^{-12} s. On peut également utiliser pour cela un *éclairement rasant de la photocathode* (angle d'incidence ~85°). A l'aide de ces artifices, on arriva à déterminer sûrement le seuil photoélectrique des différents métaux (Na, Ag, Au, etc.) et des différents semiconducteurs pour $N = 2, 3, 4, 5$ en faisant varier l'intensité lumineuse de 0,1 jusqu'à 10^3 MW. L'étude de l'influence de la polarisation et de l'angle d'incidence du rayonnement laser sur le courant de saturation a permis de constater que l'effet à plusieurs photons était dû exclusivement à la composante du champ électrique, *qui est normale à la surface de la cathode.*

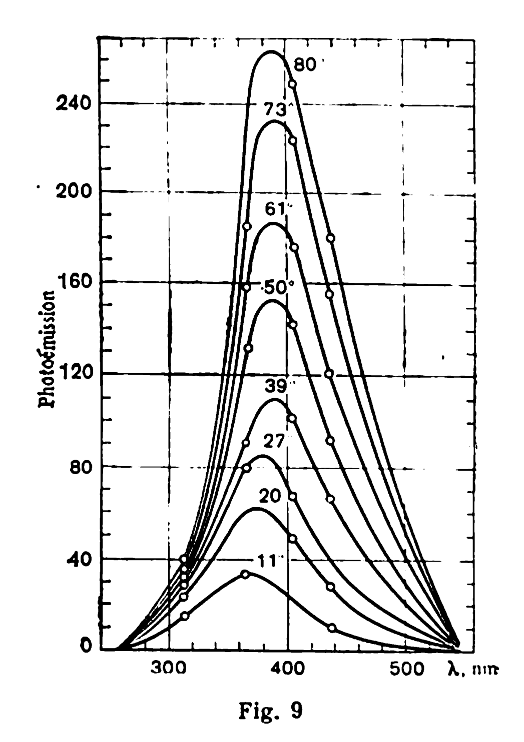


Fig. 9

10. Pour conclure ce paragraphe, examinons brièvement *l'effet photoélectrique intérieur* ainsi que les applications de l'effet photoélectrique. L'effet photoélectrique intérieur s'observe dans les semiconducteurs et les diélectriques. Sous l'action de la lumière, une partie des électrons est transférée de la bande de valence dans la bande de conduction (cf. t. III, § 100). A l'intérieur du corps, la concentration des porteurs de courant et la conductibilité électrique du corps augmentent ; c'est l'effet de *photoconduction.* La redistribution des électrons parmi les états énergétiques peut également donner lieu à une variation du *champ électrique intérieur* dans le cristal. Cela conduit à l'apparition d'une *force photoélectromotrice* au contact de deux semiconducteurs de types opposés, ou entre un métal et un semiconducteur lorsqu'ils sont éclairés. A proximité de la frontière entre les corps, il se forme une *zone de transition* qui ne laisse passer le courant que dans un seul sens, et présente donc un *effet de valve* (cf. t. III, § 100).

L'ffet photoélectrique (extérieur ou intérieur) est mis à profit dans les *dispositifs photoélectroniques* trouvant de nombreuses applications dans les sciences et les techniques (télévision, techniques aérospatiales, etc.). On utilise largement les *cellules photoémissives* à

deux électrodes, entre lesquelles on établit une différence de potentiel; l'éclairement de la cathode donne lieu à la transformation de l'énergie radiante en énergie électrique. Dans les dispositifs appelés *photorésistances*, la résistance électrique du semiconducteur diminue lors de l'éclairement. L'apparition d'une force photoélectromotrice, lors de l'éclairement d'une jonction de deux semiconducteurs différents, constitue la base de fonctionnement des *photodiodes* qui réalisent la conversion directe de l'énergie radiante en énergie électrique. Les *photomultiplicateurs* (cf. t. III, § 103), qui amplifient le courant photoélectrique initial, permettent d'enregistrer des flux lumineux extrêmement faibles, même des quanta isolés

PROBLÈMES

1. Calculer la vitesse maximale des photoélectrons sortant d'une électrode de nickel, éclairée par une lumière ultraviolette, de longueur d'onde $\lambda = 220$ nm. Pour le nickel, le travail de sortie est $A = 4{,}84$ eV.

R é p o n s e. $v_{max} = c\sqrt{\frac{2}{m_e c^2}\left(\frac{hc}{\lambda} - A\right)} = 510$ km/s.

2. Calculer le nombre N d'électrons émis par seconde par la cathode d'une cellule photoélectrique à vide, lors de son éclairement, sachant que le courant de saturation est $I = 5\cdot 10^{-10}$ A.

R é p o n s e. $N = I/e = 3{,}1\cdot 10^{9}$ s^{-1}.

3. Calculer la longueur d'onde λ du seuil photoélectrique pour le zinc, sachant que pour ce métal le travail de sortie de l'électron du zinc est $A = 3{,}74$ eV.

R é p o n s e. $\lambda = hc/A = 330$ nm.

4. L'énergie cinétique maximale des photoélectrons émis lorsqu'on éclaire une électrode de césium avec de la lumière monochromatique a été trouvée égale à $\mathscr{E}_{max} = 0{,}15$ eV. Calculer la longueur d'onde λ de la lumière utilisée, sachant que pour le césium le travail de sortie de l'électron du césium est $A = 1{,}89$ eV.

R é p o n s e. $\lambda = \frac{hc}{\mathscr{E}_{max} + A} = 600$ nm.

5. Une bille de cuivre isolée est éclairée avec de la lumière ultraviolette, de longueur d'onde $\lambda = 200$ nm. Calculer le potentiel maximal auquel peut être portée la bille, si le travail de sortie des électrons du cuivre est $A = 4{,}47$ eV.

R é p o n s e. $V = \frac{hc/\lambda - A}{e} = 1{,}04$ V.

6. Reprendre les données du problème précédent et déterminer pour quelles longueurs d'onde λ de la lumière éclairant la bille de cuivre celle-ci n'acquerra aucune charge.

R é p o n s e. $\lambda \geqslant hc/\lambda = 274$ nm.

7. Une cellule photoélectrique à vide, à cathode en molybdène, est éclairée avec une lumière monochromatique de longueur d'onde $\lambda = 250$ nm. Lorsqu'on établit, entre les électrodes, une différence de potentiel ralentissant les électrons, le courant photoélectrique diminue et s'annule lorsqu'elle devient égale à $V = 1{,}8$ V. Calculer la différence de potentiel de contact externe entre le molybdène et la substance de l'anode, sachant que pour le molybdène le travail de sortie des électrons est $A = 4{,}27$ eV.

R é p o n s e. $V_c = \frac{hc/\lambda - A}{e} - V = -1{,}40$ V.

Le signe négatif signifie que lorsqu'on met le molybdène en contact avec la substance de l'anode, le potentiel du molybdène est le plus petit des deux.

8. Calculer la vitesse maximale v des électrons atteignant l'anode d'un tube à rayons X, sachant que la longueur d'onde minimale du spectre continu du rayonnement X est $\lambda = 0{,}1$ nm.

Réponse. $v = \dfrac{h}{m_e \lambda} = 7{,}3 \cdot 10^3$ **km/s.**

§ 3. Effet Compton

1. En 1922, Arthur Compton (1892-1962) découvrit un phénomène qui confirme, tout comme l'effet photoélectrique, l'hypothèse photonique. Compton étudiait la diffusion des rayonnements X *durs* par des corps composés d'*atomes légers* (graphite, paraffine, etc.), à l'aide d'un équipement dont le schéma est représenté sur la figure 10.

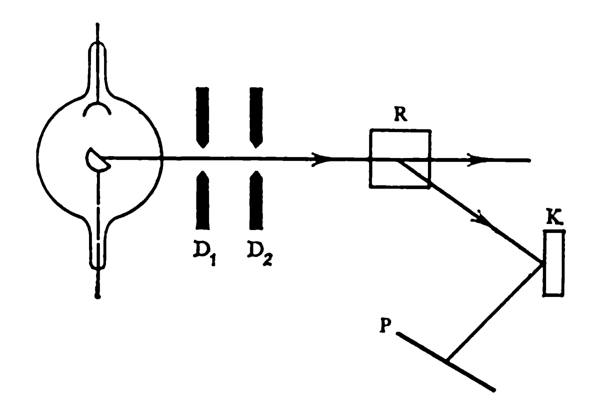


Fig. 10

La source de rayonnement X était un tube à anticathode en molybdène. Un faisceau monochromatique étroit, délimité par les diaphragmes D_1 et D_2, était diffusé par le corps R étudié. Pour pouvoir déterminer la composition spectrale de la radiation diffusée, on l'envoyait, à travers un système de diaphragmes, sur le cristal K d'un spectrographe à rayons X, et de là, dans une chambre d'ionisation ou sur une plaque photographique. Compton constata que la radiation diffusée contenait, en plus de la longueur d'onde initiale λ, des rayons d'une autre longueur d'onde $\lambda' > \lambda$ (*raie déplacée*). Le déplacement $\lambda' - \lambda$ (vers les ondes longues) résultant de la diffusion de la radiation initiale, a été appelé *déplacement de Compton*, et le phénomène a reçu le nom d'*effet Compton.* L'expérience montre que *pour tous les corps étudiés le déplacement de Compton $\lambda' - \lambda$ est indépendant de la nature des substances diffusantes et de la longueur d'onde λ de la radiation incidente.* Ce déplacement est proportionnel au carré du sinus de l'angle de diffusion ϑ divisé par deux.

La figure 11 représente les résultats des mesures dans le cas du graphite, pour différents angles de diffusion, à la longueur d'onde de la raie K du molybdène ($\lambda = 0{,}0712605$ nm). En haut on a représenté la forme de la raie de la radiation initiale (i.e. la distribution angulaire des intensités pour cette raie). En bas, on a représenté la même chose pour la radiation diffusée, pour différentes valeurs de l'angle de diffusion ϑ. On voit nettement que la raie initiale simple se dédouble par diffusion. L'élargissement des deux composantes de la raie double est déterminé par le mouvement des électrons et des atomes diffusant la radiation.

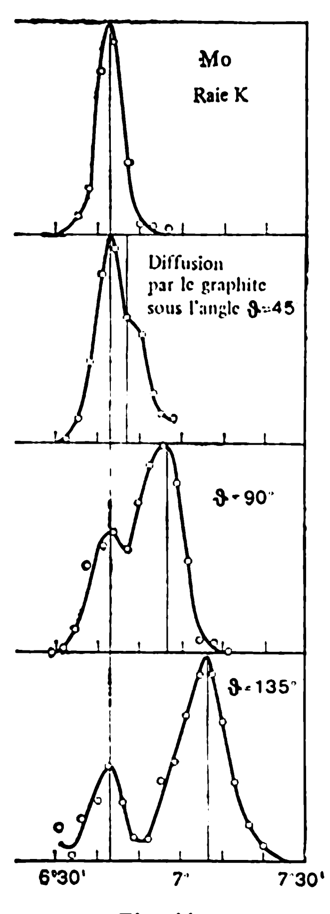


Fig. 11

2. La théorie classique fut incapable d'expliquer l'effet Compton (voir plus loin sous-paragraphe 9). On n'arriva à expliquer ce phénomène que sur la base d'une *théorie quantique* élaborée par Compton et Debye (1884-1966). Selon cette théorie, la diffusion d'un quantum X, avec changement de longueur d'onde, résulte d'une collision *unique* entre le quantum incident et un électron.

Dans les substances légères, utilisées dans les expériences de Compton, l'énergie de liaison de l'électron avec l'atome est petite devant l'énergie que lui cède le photon au cours de la collision. L'énergie que le photon cède à l'atome au cours de la collision est d'autant plus grande que l'angle de diffusion est grand. Il s'ensuit que la condition ci-dessus est d'autant mieux vérifiée que l'angle de diffusion est grand. Pour les atomes légers, on peut négliger l'énergie de liaison des électrons dans l'atome pour tous les angles de diffusion, ce qui permet de poser que tous les électrons sont libres. C'est ce que postule la théorie de Compton-Debye. On comprend alors que le déplacement de Compton $\lambda' - \lambda$ soit le même pour toutes les substances. Dans cette théorie, on admet dès le début que la substance diffusante est, en fait, composée uniquement d'*électrons libres*, ce qui implique que les particularités individuelles des substances n'importent point. Mais cette idéalisation ne vaut que pour les atomes légers et ne s'applique plus aux électrons des couches intérieures des atomes lourds. La prise en compte de l'énergie de liaison

doit conduire à des écarts à la formule simple (3.3) qu'on déduit ci-après. On décela effectivement des écarts à cette formule.

3. Analysons la collision entre un photon et un électron libre. Lors des interactions de ces deux types de particules, les *lois de la conservation de l'énergie de l'impulsion* doivent être respectées. Comme à la suite de la collision avec un photon l'électron peut acquérir une vitesse relativiste, l'étude de la collision doit être fondée sur la *mécanique relativiste.* Notons $\mathscr{E}_{ph}$ et p_{ph} l'énergie et l'impulsion du photon avant collision, et notons $\mathscr{E}'_{ph}$ et p'_{ph} les valeurs de ces grandeurs après collision. Pour l'électron, l'énergie totale et l'impulsion avant collision sont respectivement égales à $\mathscr{E}_0 = m_0c^2$ et à 0 (avant collision, l'électron était au repos) ; après collision, ces grandeurs deviennent égales à $\mathscr{E}_e$ et p_e. En appliquant les lois de la conservation de l'énergie et de l'impulsion, on obtient

$$\mathscr{E}_{ph} + \mathscr{E}_0 = \mathscr{E}'_{ph} + \mathscr{E}_e,$$

$$p_{ph} = p'_{ph} + p_e,$$

ou

$$\mathscr{E}_e = \mathscr{E}_{ph} - \mathscr{E}'_{ph} + \mathscr{E}_0,$$

$$p_e = p_{ph} - p'_{ph}.$$

On en déduit

$$\left(\frac{\mathscr{E}_e}{c}\right)^2 - p_e^2 = \frac{(\mathscr{E}_{ph} - \mathscr{E}'_{ph} + \mathscr{E}_0)^2}{c^2} - (p_{ph} - p'_{ph})^2. \tag{3.1}$$

Pour toute particule, le carré du quadrivecteur est un invarian dans l'univers de Minkowski (1864-1909), i.e.

$$\left(\frac{\mathscr{E}}{c}\right)^2 - p^2 = \text{invariant}. \tag{3.2}$$

Pour le photon, cet invariant est égal à zéro. Il s'ensuit que

$$\left(\frac{\mathscr{E}_e}{c}\right)^2 - p_e^2 = \left(\frac{\mathscr{E}_0}{c}\right)^2,$$

$$\left(\frac{\mathscr{E}_{ph}}{c}\right)^2 = p_{ph}^2,$$

$$\left(\frac{\mathscr{E}'_{ph}}{c}\right)^2 = p'^2_{ph}.$$

Compte tenu de ces résultats, la formule (3.1) donne

$$\frac{\mathscr{E}_{ph}\mathscr{E}'_{ph}}{c^2} + \frac{\mathscr{E}'_{ph}\mathscr{E}_0}{c^2} - \frac{\mathscr{E}_{ph}\mathscr{E}_0}{c^2} - p_{ph}p'_{ph} = 0,$$

ou

$$1 - \cos\vartheta = m_0c\left(\frac{1}{p'_{ph}} - \frac{1}{p_{ph}}\right),$$

ϑ étant l'angle de diffusion, i.e. l'angle que font les vecteurs p_{ph} et p'_{ph}.

Jusqu'ici, nous n'avons pas fait intervenir la nature quantique de la lumière, tout au moins sous sa forme spéciale (1.7), puisque les notions d'énergie et d'impulsion du rayonnement étaient suffisantes. Les considérations ci-dessus sont également valables en *physique classique* puisque celle-ci attribue aux radiations une énergie et une impulsion déterminées vérifiant les lois de la conservation. C'est en appliquant ces lois de la conservation que nous avons analysé l'interaction de l'électron avec une portion déterminée de radiation. Il convient maintenant de prendre en considération que cette portion de radiation est le photon d'impulsion initiale $p_{ph} = h/\lambda$ et d'impulsion finale $p'_{ph} = h/\lambda'$, λ et λ' étant les longueurs d'onde de la radiation incidente et de la radiation diffusée. On obtient

$$\lambda' - \lambda = \lambda_C (1 - \cos\vartheta) = 2\lambda_C \sin^2\frac{\vartheta}{2}, \tag{3.3}$$

où

$$\lambda_C = \frac{h}{m_e c} = 2{,}4263096\,(15)\cdot 10^{-10}\ \text{cm}. \tag{3.4}$$

Dans (3.4), m_e désigne la masse au repos de l'électron (voir fin du § 1).

La relation (3.3) établit l'indépendance du déplacement de Compton $\lambda' - \lambda$ avec la nature de la substance diffusante (pour les atomes légers). Nous avons déjà décrit ce résultat de façon qualitative. La formule (3.3) démontre aussi que le déplacement de Compton ne dépend pas de la longueur d'onde λ de la radiation incidente. Ce dernier résultat est spécifique de l'étude *relativiste*, à condition de remarquer que la masse au repos de l'un des corps (le photon) est égale à zéro.

La constante universelle λ_C, définie par la formule (3.4), est l'une des principales constantes atomiques. On la désigne sous le nom de *longueur de Compton pour l'électron*. C'est le changement de la longueur d'onde du photon résultant de sa diffusion d'un angle $\vartheta = \pi/2$ par un électron libre en position fixe. Une autre signification physique de la longueur de Compton sera donnée au § 17. Il existe une longueur de Compton pour le proton, le neutron et les autres particules élémentaires ; elle se définit également par l'expression (3.4), à condition d'y remplacer la masse de l'électron par la masse de la particule considérée. Dans les études théoriques, on préfère utiliser une autre constante universelle définie par la relation

$$\lambdabar_C = \frac{\hbar}{m_e c} = \frac{\lambda_C}{2\pi} = 3{,}861592\,(4)\cdot 10^{-11}\ \text{cm}. \tag{3.5}$$

Cette constante est également désignée sous le nom de *longueur de Compton pour l'électron* (*réduite* ou *barrée*).

On notera que dans cette théorie l'électron qui diffuse le photon est supposé être immobile. S'il était en mouvement, lors de la collision il pourrait céder son énergie cinétique au photon et s'immobiliser. Ce processus donnerait lieu à une diminution de la longueur d'onde du photon et c'est pour cela qu'on l'a dénommé *effet Compton inverse*.

4. La formule (3.3) montre que la diffusion des photons par les électrons libres immobiles doit toujours donner lieu à un accroissement de la longueur d'onde par effet Compton. Quelle est donc l'origine de la raie non déplacée (voir fig. 11)? Elle est tout simplement due aux électrons liés. Dans ce cas, les photons sont diffusés par des *atomes* dont la masse peut être considérée comme infiniment grande. Selon les formules (3.3) et (3.4), pour un atome infiniment lourd la longueur de Compton, ainsi que le déplacement $\lambda' - \lambda$, deviennent nuls. Le photon incident cède son impulsion à l'atome, mais son énergie ne change pas. A mesure qu'augmente le numéro atomique, le nombre relatif d'électrons liés augmente, et par suite l'intensité de la composante non déplacée augmente en comparaison de celle de la composante déplacée.

Il reste à préciser la variation du rapport des intensités de ces composantes en fonction de l'angle de diffusion. Nous avons signalé ci-dessus que les conditions permettant de considérer un électron comme étant libre sont d'autant mieux remplies que l'angle de diffusion est grand. Il s'ensuit que lorsque l'angle de diffusion augmente, la proportion relative d'électrons libres s'accroît, ce qui entraîne un accroissement du rapport des intensités des composantes déplacée et non déplacée.

Plus l'énergie du photon est grande, moins se manifeste la liaison des électrons avec les atomes. C'est pour cela qu'on doit utiliser un *rayonnement* X *dur* pour observer l'effet Compton. Cependant, lorsque l'énergie du photon devient supérieure au double de l'énergie de l'électron au repos $2m_ec^2$, l'interaction des photons avec la substance commence à donner lieu à la *formation de paires*, autrement dit, le photon commence à se transformer en une paire électron-positon. Ce processus s'intensifie à mesure qu'augmente l'énergie des photons, et pour des énergies notablement supérieures à $2m_ec^2$, il supplante l'effet Compton. Pour la lumière visible, l'énergie de liaison des électrons dans les atomes est plus grande que l'énergie des photons. C'est pour cela qu'on n'observe pas l'effet Compton dans la partie visible du spectre.

Notons que la diffusion des photons par les atomes est *cohérente*, et leur diffusion par les électrons libres est *incohérente*. Ce dernier résultat est évident ; puisque les électrons libres et leurs mouvements sont indépendants, la diffusion des photons est également indépendante. D'autre part, les vibrations des électrons liés, déterminées par l'onde incidente, sont corrélées. Par conséquent, les ondes dif-

fusées par les électrons liés peuvent interférer entre elles ainsi qu'avec l'onde incidente. C'est précisément des interférences de ce genre qui se produisent lorsque les rayons X traversent des cristaux, et qui sont alors définies par les conditions de Laue et de Wulff-Bragg (voir t. IV, § 61).

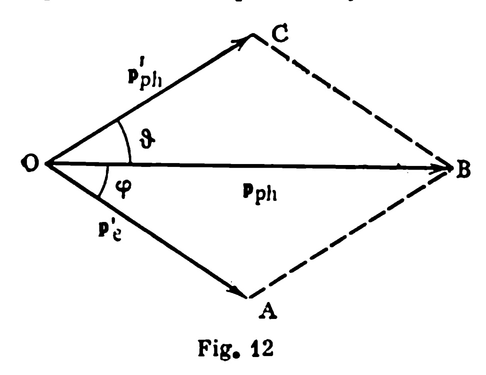


Fig. 12

5. Lorsqu'un photon X est diffusé par un électron libre, celui-ci éprouve un mouvement de recul, comme indiqué par le parallélogramme sur la figure 12. On en tire les relations

$$p'_e \cos\varphi + p'_{ph} \cos\vartheta = p_{ph},$$

$$p'_e \sin\varphi = p'_{ph} \sin\vartheta.$$

En éliminant p'_e entre ces équations et en remarquant que $p_{ph}/p'_{ph} = \lambda'/\lambda$, on obtient

$$\operatorname{tg}\varphi = \frac{\sin\vartheta}{\lambda'/\lambda - \cos\vartheta}.$$

La formule (3.3) permet de trouver λ'/λ. Après quelques réarrangements, on trouve

$$\operatorname{tg}\varphi = \frac{\operatorname{cotg}(\vartheta/2)}{1+\lambda_C/\lambda} = \frac{\operatorname{cotg}(\vartheta/2)}{1+h\nu/(m_e c^2)}. \tag{3.6}$$

Cette formule détermine la direction de sortie de l'électron de recul. Le calcul de l'énergie et de l'impulsion des électrons de recul ne présente aucune difficulté (voir problème 2 à la fin de ce paragraphe).

6. Le processus illustré par la figure 12 s'observe dans la chambre de Wilson (1869-1959). Rappelons que dans une chambre de Wilson on produit de la vapeur d'eau sursaturée, à l'aide d'une détente adiabatique rapide (voir t. II, § 119). Lorsqu'une particule chargée traverse la chambre de Wilson, elle ionise l'air ambiant. Les ions ainsi produits attirent les gouttelettes de brouillard, ce qui visualise la *trace* de la particule chargée. A l'aide de la figure 12, on peut donc déterminer la direction dans laquelle se déplacera un électron de recul dans la chambre de Wilson. L'énergie et l'impulsion de l'électron de recul dépendent de la longueur de son trajet, i.e. de la longueur de la trace qu'il laisse dans la chambre de Wilson. On peut donc considérer que l'impulsion p'_e de l'électron de recul est connue.

Il reste à voir comment on fait pour déterminer les longueurs des segments *OB* et *OC*, représentant respectivement les impulsions du photon incident et du photon diffusé. N'étant pas chargés, les photons ne laissent aucune trace de leur passage dans la chambre de Wilson. L'origine du segment *OB* est marquée par le point où commence la trace de l'électron de recul, et sa longueur est déterminée

par la valeur de l'impulsion du photon incident. Ces données sont suffisantes pour déterminer complètement le segment OB. Il reste à déterminer le segment OC; pour ce faire, il suffit de connaître la position du point C. La position du point C sera connue si le photon diffusé subira une diffusion de Compton qui fera apparaître un nouvel électron de recul. Dans ce dernier cas, les trois segments OA, OB et OC seront connus, et on constatera alors que OB est bien la diagonale du parallélogramme construit sur les segments OA et OC, conformément à la théorie quantique de l'effet Compton. On peut vérifier notamment que les angles de diffusion du photon et de l'électron de recul satisfont à la relation (3.6). Sur la base d'une seule expérience on ne peut pas affirmer que la nouvelle trace est laissée non par un électron aléatoire quelconque qui n'a aucun rapport au phénomène étudié mais par un *nouveau électron de recul* dont il s'agissait plus haut. On ne peut l'affirmer qu'avec une *certaine probabilité*, à la suite d'un traitement statistique des résultats d'un très grand nombre d'expériences. C'est une telle statistique qui a permis de justifier les hypothèses de base de la théorie photonique de l'effet Compton. En découle notamment la preuve que *les lois de la conservation de l'énergie et de l'impulsion sont vérifiées dans chaque acte élémentaire de diffusion des photons par les électrons.*

8. D. V. Skobeltsyne (né en 1892) a observé des électrons de recul dans une chambre de Wilson, soumise à l'action d'un intense champ magnétique lorsqu'on l'irradiait avec les rayons γ émis par une source radioactive. Ces électrons de recul avaient des trajectoires circulaires de rayon

$$R = \frac{cp'_e}{eB}, \tag{3.7}$$

où B dénote l'intensité du champ magnétique. Ayant mesuré le rayon des trajectoires, l'application de cette formule permet de calculer l'impulsion, puis, connaissant l'impulsion, calculer l'énergie de l'électron par la formule

$$\mathscr{E}_e^2 = (p'_e c)^2 + (m_0 c^2)^2. \tag{3.8}$$

On constata alors qu'il s'agissait bien d'*électrons relativistes*, ce qui imposait la mise en œuvre de la théorie de la relativité.

9. Partant de cette dernière assertion, il est tout indiqué de démontrer qu'aucune interprétation classique de l'effet Compton n'est possible. Puisqu'on peut admettre que les électrons des atomes légers sont libres, ceux-ci ne peuvent posséder des fréquences propres. Examinons, dans le cadre de la théorie classique, le comportement de ces électrons lorsqu'ils sont soumis au champ d'une onde électromagnétique monochromatique. Pour commencer, négligeons l'action du champ magnétique. Dans cette approximation et en négligeant encore l'amortissement, l'électron se met à exécuter des os-

cillations, dont la fréquence est égale à celle de l'onde électromagnétique. L'électron oscillant réémet des ondes électromagnétiques qui constituent la radiation diffusée. Ainsi, du point de vue classique, la fréquence de la radiation diffusée doit coïncider avec celle de la radiation incidente.

Ce résultat a été obtenu en négligeant l'action du champ magnétique. Si on tient compte du champ magnétique B, et en négligeant toujours les amortissements, l'équation du mouvement de l'électron libre s'écrira

$$\dot{p} = -e\left(E + \frac{1}{c}[vB]\right).$$

Dans l'approximation non relativiste, c'est une équation différentielle linéaire à coefficients variables. Il s'ensuit que dans le cas d'oscillations forcées des électrons, on peut voir apparaître des fréquences différant de la fréquence de l'onde incidente. Il est bien évident que cette possibilité ne sera réalisée que si la vitesse v de l'électron n'est pas négligeable devant la vitesse de la lumière. Cette éventualité devient notable aux vitesses relativistes de l'électron surtout lorsque son mouvement est affecté par des effets non linéaires. Or, dans l'effet Compton, apparaissent justement des électrons de recul relativistes, comme le démontrent les expériences avec la chambre de Wilson. On pourrait penser que par diffusion sur des électrons libres on pourrait observer un changement de fréquence de l'onde électromagnétique.

Mais même si ce changement de fréquence était conforme à la loi de variation avec l'angle de diffusion, l'espoir d'expliquer l'effet resterait vain. En effet, il suffit de se demander pour quelle raison pourraient apparaître des électrons de recul relativistes dans le cas d'une onde électromagnétique continue. (L'expérience montre que, tout comme dans le cas de l'effet photoélectrique, les électrons de recul dus à un éclairement par rayons gamma apparaissent sans aucun retard!) Il est inconcevable qu'une onde électromagnétique continue puisse se contracter en un paquet d'énergie qui agirait sur un électron isolé! On doit donc reconnaître que les résultats expérimentaux nous imposent d'adopter le point de vue quantique, selon lequel la radiation diffusée par effet Compton et les électrons de recul résultent tous deux de l'*action isolée* d'un quantum de lumière sur un électron isolé.

PROBLÈMES

1. Après une diffusion de Compton par un électron libre immobile, la longueur d'onde d'un photon d'énergie $\mathscr{E}_{\text{ph}}$ augmente de α fois. Calculer l'énergie cinétique $\mathscr{E}_e$ de l'électron de recul.

Réponse. $\mathscr{E}_e = \frac{\alpha - 1}{\alpha}\mathscr{E}_{\text{ph}}$.

2. Un photon de rayonnement X d'énergie $\mathscr{E}_{\text{hp}}$ dévie d'un angle ϑ de sa direction initiale à la suite d'une diffusion de Compton par un électron libre

immobile. Calculer l'énergie cinétique $\mathscr{E}_{\text{cin}}$ et l'impulsion p'_e de l'électron de recul. En se fondant sur une construction géométrique, démontrer qu'en valeur absolue l'impulsion de l'électron sera plus grande que l'impulsion du photon incident, si ce dernier dévie de sa direction initiale d'un angle $\vartheta = \pi/2$.

R é p o n s e. $\mathscr{E}_{\text{cin}} = \frac{hc}{\lambda}\left(1 - \frac{\lambda}{\lambda + 2\lambda_C \sin^2(\vartheta/2)}\right)$,

$$(p'_e c)^2 = \mathscr{E}^2_{\text{cin}} + 2m_e c^2 \mathscr{E}_{\text{cin}}.$$

3. Un photon de longueur d'onde λ a été diffusé par un électron libre en mouvement. A la suite de cet acte de diffusion, l'électron a été immobilisé, et le photon a été dévié d'un angle ϑ par rapport à sa direction initiale. Calculer le changement de longueur d'onde $\lambda' - \lambda$ correspondant. Ramener ce problème à celui de la diffusion d'un photon par un électron immobile.

R é p o n s e. $\lambda' - \lambda = -\frac{2h}{m_e c}\sin^2(\vartheta/2)$. Par suite de la diffusion la longueur d'onde du photon diminue.

4. Un quantum gamma d'énergie $h\nu$ est diffusé par un électron immobile. Déterminer la direction du quantum diffusé qui, par interaction avec une substance, donnerait naissance à une paire électron-positon. Calculer la valeur limite de l'angle de diffusion ϑ_{lim} pour laquelle ce processus est encore réalisable dans le cas de quanta γ de très grandes énergies.

R é p o n s e. Un quantum gamma diffusé ne peut donner naissance à une paire électron-positon que si la diffusion a lieu dans un cône d'angle au sommet défini par

$$\cos\vartheta > \frac{1}{2} + \frac{m_e c^2}{h\nu}.$$

Si la direction de la diffusion se situe en dehors de ce cône, aucune paire électron-positon ne peut être générée. Pour des quanta de très grandes énergies ($h\nu \to \infty$) l'angle de diffusion ne doit pas dépasser $\vartheta_{\text{lim}} = 60°$ pour que la génération des paires soit encore possible.

5. Démontrer qu'aussi grande que soit l'énergie d'un photon, celui-ci ne pourra donner naissance à une paire électron-positon, si à la suite de sa diffusion par un électron immobile l'angle de diffusion est exactement égal à $\vartheta = \pi$.

6. Soit ν la fréquence d'un photon avant qu'il ait éprouvé une diffusion de Compton par un électron dans un référentiel, où l'électron est au repos après sa collision avec le photon, et soit ν' la fréquence du photon après collision avec un électron dans un référentiel, dans lequel l'électron était initialement au repos (i.e. dans un référentiel lié au laboratoire). Démontrer que $\nu = \nu'$ (théorème de réciprocité).

S o l u t i o n. Notons par les lettres majuscules $\mathscr{P}$ et $\mathscr{K}$ les impulsions quadridimensionnelles $\left(p, \frac{\mathscr{E}_e}{c}\right)$ de l'électron et du photon avant collision. Après collision, ces mêmes grandeurs seront notées par les mêmes lettres mais affectées de primes. Dans le cas d'une collision élastique, il y a conservation de l'impulsion quadridimensionnelle du système photon-électron:

$$\mathscr{P} + \mathscr{K} = \mathscr{P}' + \mathscr{K}'.$$

En récrivant cette relation sous la forme

$$\mathscr{P} - \mathscr{K}' = \mathscr{P}' - \mathscr{K},$$

et l'élevant au carré, on obtient

$$\mathscr{P}^2 + \mathscr{K}'^2 - 2\mathscr{P}\mathscr{K}' = \mathscr{P}'^2 + \mathscr{K}^2 - 2\mathscr{P}'\mathscr{K}.$$

Dans l'univers de Minkowski *), le carré de la quadri-impulsion de la particule est invariant. Pour un électron on a

$$\mathcal{P}^2 = p^2 - \frac{\mathcal{E}_c^2}{c^2} = p'^2 - \frac{\mathcal{E}_c'^2}{c^2} = -m_0^2 c^2,$$

où m_0 dénote la masse au repos de l'électron. Pour le photon le même invariant est nul: $\mathcal{K}^2 = 0$ (puisque la masse au repos du photon est nulle). Ainsi

$$\mathcal{P}\mathcal{K}' = \mathcal{P}'\mathcal{K}.$$

Dans le référentiel où, avant collision, l'électron se trouvait au repos, on a $\mathcal{P} = (0, m_0 c)$, $\mathcal{K}' = \left(p'_{\text{ph}}, \frac{h\nu'}{c}\right)$, de sorte que $\mathcal{P}\mathcal{K}' = -m_0 h\nu'$. Dans le référentiel où l'électron était au repos après avoir subi la collision on a : $\mathcal{P}' = (0, m_0 c)$, $\mathcal{K} = \left(p_{\text{ph}}, \frac{h\nu}{c}\right)$, d'où $\mathcal{P}'\mathcal{K} = -m_0 h\nu$. On en déduit $\nu = \nu'$, C.Q.F.D.

7. Démontrer que dans un référentiel, dans lequel l'électron se trouve au repos après avoir subi une collision avec un photon, la fréquence de celle-ci, après une diffusion de Compton, est égale à la fréquence des photons incidents, mesurée dans un référentiel de laboratoire.

8. Démontrer que lors d'une diffusion de Compton dans un référentiel constamment lié à l'électron, les fréquences du photon incident et du photon diffusé ont égales, tandis que les impulsions sont égales mais de sens opposés.

§ 4. Application de la théorie photonique à l'étude de l'effet Doppler, résultant du mouvement d'une source lumineuse dans le vide

1. L'effet photoélectrique et l'effet Compton sont deux phéno mènes typiquement *quantiques* ne pouvant être interprétés par la théorie classique. Dans ce paragraphe, ainsi que les trois paragraphes suivants, on considérera des phénomènes pouvant être également bien interprétés par la théorie classique et par la théorie quantique, avec des résultats concordants. Examinons, pour commencer, l'effet Doppler (1803-1853) dans l'approximation non relativiste. Choisissons un référentiel d'inertie dans lequel la source lumineuse de masse M se déplace dans le vide avec la vitesse v. L'énergie de la source se compose de l'énergie cinétique $1/2\ Mv^2$ et de l'énergie interne $\mathcal{E}$ des atomes excités. Lors de l'émission de lumière, l'énergie interne de la source varie. Notons $\mathcal{E}$ sa valeur initiale et $\mathcal{E}'$ sa valeur finale. D'autre part, du fait de la pression de radiation, la source subit un recul, et sa vitesse prend un accroissement $(v' - v)$. En vertu des lois de la conservation de l'énergie et de l'impulsion, on a

$$\tfrac{1}{2} Mv^2 + \mathcal{E} = \tfrac{1}{2} Mv'^2 + \mathcal{E}' + \mathcal{E}_{\text{rad}}, \tag{4.1}$$

$$Mv = Mv' + p_{\text{rad}}, \tag{4.2}$$

*) Rappelons que dans l'univers de Minkowski le carré d'un quadrivecteur est égal à la différence des carrés des composantes spatiale et temporelle. Dans le produit scalaire des quadrivecteurs, le produit des composantes temporelles figure aussi avec le signe moins.

où $\mathscr{E}_{rad}$ et p_{rad} sont respectivement l'énergie et l'impulsion de la radiation dans le référentiel d'inertie considéré.

Elevons au carré la seconde équation, divisons le résultat obtenu par $2M$ et soustrayons de la première équation. On obtient alors

$$\mathscr{E}-\mathscr{E}'=\mathscr{E}_{rad}-v'p_{rad}-\frac{p_{rad}^2}{2M},$$

et en vertu de (4.2)

$$\mathscr{E}-\mathscr{E}'=\mathscr{E}_{rad}-vp_{rad}+\frac{p_{rad}^2}{2M}.$$

Si la masse M de la source est grande, on peut négliger le dernier terme de cette équation Dans cette approximation on a

$$\mathscr{E}-\mathscr{E}'=\mathscr{E}_{rad}-vp_{rad}. \tag{4.3}$$

On peut encore éliminer de cette expression l'impulsion de la radiation. Comme la radiation est un *objet* essentiellement *relativiste*, son impulsion se laisse exprimer en fonction de l'énergie à l'aide de la relation $p_{rad} = \mathscr{E}_{rad}/c$. En substituant cette dernière expression dans (4.3), on obtient

$$\mathscr{E}-\mathscr{E}'=\mathscr{E}_{rad}\left(1-\frac{v}{c}\cos\vartheta\right), \tag{4.4}$$

où ϑ est l'angle entre la direction de mouvement du corps émetteur et la direction de la radiation (c'est donc l'angle entre v et p).

La relation (4.4) a été établie *sans faire appel à aucune conception quantique*. On n'a utilisé que les lois de la conservation de l'énergie et de l'impulsion, qui sont également valables en physique classique et en physique quantique. Admettons maintenant que la radiation s'effectue par émission de photons. Supposons qu'un seul photon a été émis. En notant ν' la fréquence de ce photon, repérée dans le référentiel dans lequel se meut l'émetteur (c'est le référentiel d'inertie que nous avons choisi), on aura $\mathscr{E}_{rad} = h\nu'$. Dans le référentiel par rapport auquel l'émetteur est au repos, nous dénoterons la fréquence du photon par ν. En considérant le processus d'émission du photon dans ce dernier référentiel, on devra écrire $\mathscr{E} - \mathscr{E}' = h\nu$, puisque dans ce cas la variation d'énergie interne de l'émetteur résulte uniquement de l'émission d'un quantum de lumière. En portant ces expressions dans (4.4) et en divisant par h, on obtient

$$\nu'=\frac{\nu}{1-\frac{v}{c}\cos\vartheta}. \tag{4.5}$$

Cette formule n'est exacte qu'au *premier ordre*, i.e. aux termes du premier degré en v/c près, et ce pour la raison qu'elle a été établie dans le cadre de la mécanique non relativiste. Au même ordre de pré-

cision on peut remplacer (4.5) par la formule suivante:

$$\nu' = \nu\left(1 + \frac{v}{c}\cos\vartheta\right). \tag{4.6}$$

2. Etablissons maintenant pour l'effet Doppler une *formule relativiste*, valable pour toutes les vitesses v. Dans ce cas, il n'est plus nécessaire de subdiviser l'énergie totale (relativiste) du corps en énergie interne et en énergie cinétique. Les symboles $\mathcal{E}$ et $\mathcal{E}'$ désigneront désormais l'*énergie totale* du corps dans le référentiel d'inertie considéré avant et après émission de lumière. Notons $\mathcal{E}_0$ et $\mathcal{E}_0'$ les énergies correspondantes pour le corps au repos avant et après émission. Au début, on utilisera à nouveau les énergies et les impulsions, et on n'introduira les conceptions quantiques qu'à la dernière étape des calculs.

Pour le processus d'émission de la lumière, il est commode d'écrire les lois de la conservation de l'énergie et de l'impulsion sous la forme suivante:

$$\mathcal{E}' = \mathcal{E} - \mathcal{E}_{\text{rad}}, \quad p' = p - p_{\text{rad}}. \tag{4.7}$$

Elevons ces deux égalités au carré après avoir multiplié la seconde par c; soustrayons membre à membre les résultats obtenus. En tenant compte des relations

$$\mathcal{E}^2 - (pc)^2 = \mathcal{E}_0^2, \quad \mathcal{E}'^2 - (p'c)^2 = \mathcal{E}_0'^2, \quad \mathcal{E}_{\text{rad}} = cp_{\text{rad}},$$

on obtient

$$\mathcal{E}_0'^2 = \mathcal{E}_0^2 - 2\mathcal{E}\mathcal{E}_{\text{rad}} + 2c^2 p p_{\text{rad}}.$$

En remarquant que

$$p = mv = \frac{\mathcal{E}}{c^2}v = \frac{\mathcal{E}\beta}{c},$$

où $\beta = v/c$, l'équation précédente devient

$$\mathcal{E}_0^2 - \mathcal{E}_0'^2 = 2\mathcal{E}\mathcal{E}_{\text{rad}}(1 - \beta\cos\vartheta), \tag{4.8}$$

ou

$$\mathcal{E}_0 - \mathcal{E}_0' = \frac{2\mathcal{E}}{\mathcal{E}_0 + \mathcal{E}_0'}\mathcal{E}_{\text{rad}}(1 - \beta\cos\vartheta).$$

De nouveau ϑ dénote ici l'angle que fait la direction de mouvement du corps émetteur avec la direction de propagation de la radiation. Si la masse du corps est grande, on peut poser $\mathcal{E}_0' = \mathcal{E}_0$ dans le dénominateur de la dernière relation. En remarquant encore que $\mathcal{E} = \mathcal{E}_0/\sqrt{1 - \beta^2}$, on arrive à la relation

$$(\mathcal{E}_0 - \mathcal{E}_0')\sqrt{1 - \beta^2} = \mathcal{E}_{\text{rad}}(1 - \beta\cos\vartheta). \tag{4.9}$$

Utilisons maintenant les relations quantiques $\mathcal{E}_0 - \mathcal{E}_0' = h\nu$ et $\mathcal{E}_{\text{rad}} = h\nu'$, qui sont exactement les mêmes que dans le cas non

relativiste. On obtient alors la *formule relativiste* de l'effet Doppler :

$$\nu' = \nu \frac{\sqrt{1-\beta^2}}{1-\beta\cos\vartheta}. \tag{4.10}$$

3. Les deux formules non relativiste (4.5) et relativiste (4.10), qu'on a établies en théorie photonique pour décrire l'effet Doppler, coïncident avec les formules correspondantes de la théorie ondulatoire classique de la lumière (voir t. IV, §§ 107, 108). La raison profonde de cette coïncidence réside en ce qu'en théorie photonique, les deux formules se déduisent des relations (4.3) et (4.9). Comme ces relations sont établies uniquement à partir des lois de la conservation de l'énergie et de l'impulsion dans les processus d'émission de la lumière, elles restent valables, que la lumière soit conçue comme des ondes continues ou comme composée de photons isolés. La seule hypothèse de caractère quantique, qu'on a invoquée pour établir la formule de la fréquence de la lumière émise, était que l'énergie des photons dépendait de façon univoque de leur fréquence. En théorie photonique, la relation entre ces grandeurs doit être de la forme spéciale : l'énergie du photon est proportionnelle à la fréquence de la lumière. Quant au coefficient de proportionnalité, i.e. la constante de Planck h, il s'élimine entre les formules définitives définissant la fréquence des vibrations. Pour s'assurer que la relation entre $\mathscr{E}$ et ν doit bien être de la forme indiquée, il suffit de remarquer qu'en appliquant la transformation de Lorentz, les grandeurs $\mathscr{E}$ et ν se transforment de la même façon. C'est pour cela que la relation $\mathscr{E} = h\nu$ présente une invariance relativiste.

Nous reprendrons l'étude de l'effet Doppler au § 6 et on considérera le cas où la source lumineuse se meut non plus dans le vide, mais dans un milieu d'indice de réfraction n.

§ 5. Réflexion et réfraction de la lumière en théorie photonique. Comportement des photons dans les milieux

1. Au XIX[e] siècle, on était persuadé que l'expérience de Foucault avait définitivement évincé la théorie corpusculaire de la lumière. Cette expérience démontra en effet que dans l'eau la vitesse de la lumière est plus petite que dans le vide. Or, selon la théorie ondulatoire, ce rapport des vitesses ne concerne que la *vitesse de phase* de la lumière. Si la lumière était effectivement de nature ondulatoire, on ne pourrait mesurer, dans l'expérience de Foucault, que la *vitesse de groupe* (cf. t. IV, § 108). Par conséquent, en toute logique, l'argumentation avancée par les physiciens du XIX[e] siècle aurait dû indiquer que dans l'expérience de Foucault la vitesse de la lumière était mesurée dans des milieux pratiquement non dispersifs. Dans ce cas, les vitesses de phase et de groupe coïncident. Leur conclusion concernant la déficience de la théorie corpusculaire est correcte, à cela

près qu'elle concerne non pas toute la théorie corpusculaire, mais la théorie corpusculaire *newtonienne*. On doit se demander si la théorie photonique d'Einstein est adéquate.

Dans la théorie corpusculaire de Newton, on expliquait la réflexion et la réfraction de la lumière sur la base de la mécanique classique de Newton. Mais cette dernière ne peut s'appliquer aux photons parce que la nature des photons est empreinte de *dualité*. Lors de leur propagation, les photons se comportent en *ondes*, et lors de leurs interactions avec les substances, ils manifestent des *propriétés de corpuscules*. Comme la réflexion et la réfraction sont des cas particuliers de la propagation de la lumière, l'application de la théorie photonique à ces deux phénomènes doit conduire aux mêmes résultats que la théorie ondulatoire classique.

2. Comme cette question mérite une étude plus détaillée, on l'examinera de plus près. Comme dans le cas du vide, on peut concevoir que dans un milieu la lumière se propage sous forme de photons. On admettra que le milieu est *immobile*; autrement dit, on envisagera le cas où les phénomènes se produisent dans un référentiel par rapport auquel le milieu est fixe. Il s'agit donc d'un référentiel dit *privilégié*. Il s'agit de déterminer l'énergie et l'impulsion des photons dans le milieu. On notera tout d'abord que *l'énergie et l'impulsion d'un photon se composent des énergies et des impulsions du champ électromagnétique et du milieu*. Dans le cas qui nous concerne, il s'agira de la somme de ces grandeurs, i.e. de l'énergie *totale* et de l'impulsion *totale* liées à la propagation d'un photon dans le milieu. On posera que le milieu est isotrope et on caractérisera ses propriétés par l'indice de réfraction $n(\omega)$.

Selon les conceptions de la théorie photonique, *le nombre total de photons ne change pas* lors de la réflexion et de la réfraction. *Si la frontière est fixe, les énergies du photon réfléchi et du photon transmis ne changent pas.* Nous allons démontrer cette assertion d'abord pour la réflexion, en supposant que le photon tombe sur la frontière en venant du vide. L'énergie du photon incident est $\mathscr{E}_e = \hbar\omega$ et dépend donc univoquement de sa fréquence ω. La même relation entre l'énergie et la fréquence doit subsister pour le photon réfléchi puisque ce dernier se propage dans le vide, tout comme le photon incident. Or, si la réflexion se produit sur une frontière immobile, la fréquence ω du photon ne change pas, sinon on observerait un changement de couleur du rayon réfléchi, chose qui ne se produit pas. Il s'ensuit que l'énergie du photon réfléchi $\mathscr{E}_r$ est égale à $\hbar\omega$, i.e. $\mathscr{E}_r = \mathscr{E}_e$.

3. La même assertion n'est pas aussi évidente pour le photon transmis. Pour la démontrer considérons un faisceau plan-parallèle composé de N_e photons monochromatiques qui, venant du vide, tombent sous n'importe quel angle d'incidence sur une frontière immobile du milieu. Soient N_r le nombre de photons réfléchis et N_d le nombre de photons transmis et soient $\mathscr{E}_r$ et $\mathscr{E}_d$ les énergies corres-

pondantes. En vertu de la loi de la conservation de l'énergie

$$N_e\mathscr{E}_e = N_r\mathscr{E}_r + N_d\mathscr{E}_d. \tag{5.1}$$

La condition de la conservation du nombre total de photons est

$$N_e = N_r + N_d. \tag{5.2}$$

Il a été démontré ci-dessus que $\mathscr{E}_r = \mathscr{E}_e$. Il s'ensuit que $\mathscr{E}_d = \mathscr{E}_e$, C.Q.F.D. On en conclut que pour les photons se propageant dans un milieu, on a toujours $\mathscr{E} = \hbar\omega$.

Considérons maintenant la question de l'impulsion des photons se propageant dans un milieu. Tout comme pour l'énergie, on peut écrire

$$N_e p_e = N_r p_r + N_d p_d, \quad N_e = N_r + N_d, \tag{5.3}$$

où les indices e, r, d dénotent les photons incidents, réfléchis et transmis. Examinons d'abord comment se comportent les composantes tangentielles des impulsions. Dirigeons l'axe X le long de la frontière et l'axe Z perpendiculairement à celle-ci. Quelle que soit la théorie, elle doit établir une loi de réflexion selon laquelle $p_{ex} = p_{rx}$. Mais on obtiendra alors pour p_{ex}, p_{rx} et p_{dx} exactement le même système d'équations que pour les énergies correspondantes. On tire de ce système

$$p_{ex} = p_{rx} = p_{dx}, \tag{5.4}$$

ce qui signifie que les composantes tangentielles des impulsions des trois photons sont égales.

Les composantes normales se comportent tout autrement. On a toujours

$$N_e p_{ez} = N_r p_{rz} + N_d p_{dz}.$$

Mais lorsqu'un photon incident s'approche du milieu, le photon réfléchi s'en éloigne, de sorte que $p_{rz} = -p_{ez}$. On obtient ainsi

$$(N_e + N_r)\,p_{ez} = N_d p_{dz},$$

ou

$$(N_e + N_r)\,p_{ez} = (N_e - N_r)\,p_{dz},$$

ou enfin

$$p_d = \frac{N_e + N_r}{N_e - N_r}\,\frac{\cos\varphi}{\cos\psi}\,p_e,$$

où φ est l'angle d'incidence et ψ l'angle de réfraction.

Il est bien évident que la forme de la relation entre p_d et p_e est déterminée par les propriétés du milieu et ne peut dépendre du type de polarisation de la lumière incidente. Cette remarque permet de simplifier les calculs ultérieurs. Supposons que la lumière incidente soit polarisée perpendiculairement au plan d'incidence. Dans ce cas, les vecteurs électriques de tous les photons seront colinéaires et

l'application de la formule de Fresnel (cf. t. IV, § 65) fournit

$$\frac{N_r}{N_e} = \frac{n \cos \psi - \cos \varphi}{n \cos \psi + \cos \varphi}.$$

En effet, comme dans une onde monochromatique tous les photons sont cohérents, le rapport N_r/N_e définit le rapport des *amplitudes* et non pas le rapport des intensités de l'onde réfléchie et de l'onde incidente. La détermination du signe dans la formule ci-dessus est fondée sur le fait que les nombres de photons N_r et N_e sont essentiellement positifs. On obtient en définitive

$$p_d = np_e. \tag{5.5}$$

4. En résumé, on constate que l'énergie et l'impulsion des photons dans un milieu ou dans le vide sont données par les formules

$$\mathscr{E} = \hbar\omega, \quad \mathbf{p} = \hbar\mathbf{k}. \tag{5.6}$$

La seule différence réside en ce que dans un milieu le nombre d'onde est

$$k = n\frac{\omega}{c}, \tag{5.7}$$

i.e. n fois plus grand que dans le vide. En introduisant la vitesse de phase $v = c/n$ et la longueur d'onde $\lambda = 2\pi v/\omega$ correspondant au milieu considéré, on obtient la formule

$$k = \frac{\omega}{v} = \frac{2\pi}{\lambda}, \tag{5.8}$$

qui également valable pour le vide et les substances, puisque dans le vide $v = c$. Dans un milieu, la relation entre l'énergie et l'impulsion d'un photon est alors de la forme

$$\mathscr{E} = pv = pc/n. \tag{5.9}$$

Il va de soi que cette relation ne s'applique qu'à un référentiel dans lequel le milieu est au repos.

Il résulte de tout cet exposé que la théorie photonique de la réfraction est analogue à la théorie newtonienne. Dans ces deux théories, on définit l'angle de réfraction par application de la loi de la conservation de l'impulsion, plus exactement de la conservation de sa composante tangentielle. La différence entre ces deux théories consiste en ce que, dans la théorie de Newton, cette loi est rattachée à la direction de la force normale à la surface de séparation et appliquée au corpuscule de lumière ; cette approche est dénuée de sens dans le cas d'un photon. La différence quantitative entre les résultats des deux théories tient à ce que *l'impulsion du corpuscule newtonien est proportionnelle à sa vitesse v, tandis que l'impulsion du photon dans un milieu est inversement proportionnelle à sa vitesse.* C'est pour cela

que la théorie photonique fournit pour l'indice de réfraction une formule identique à celle que fournit la théorie ondulatoire classique.

La conception « photons dans un milieu » s'applique non seulement à la réfraction de la lumière, mais à de nombreux autres phénomènes dont il sera question plus loin.

5. En guise de conclusion, il importe de faire une remarque de principe, relative à l'*interprétation physique* du dualisme onde-corpuscule. Il a été déjà question de ce que le nombre total de photons ne change pas lors de la réflexion ou de la réfraction de la lumière. Si un seul photon tombe sur la surface de séparation, après incidence il sera décelé soit sous forme de photon réfléchi, soit sous forme de photon transmis. Le photon réfléchi possède une impulsion égale en module, mais de sens différent que l'impulsion du photon incident. L'impulsion du photon transmis change en direction et en module. On doit se demander si la loi de la conservation de l'impulsion ne serait pas violée dans cet *acte élémentaire* concernant un seul photon, et si cette loi ne serait pas une loi statistique, comme on l'admettait implicitement dans les sous-paragraphes 2 et 3. La théorie quantique moderne répond négativement à ces questions.

Le phénomène d'interférences des photons conduit à conclure que lorsqu'un photon tombe sur une surface de séparation, il apparaît un *nouvel état* que la théorie classique ne saurait décrire. On caractérise cet état particulier en disant que *le photon se trouve partiellement dans un état de réflexion et partiellement dans un état de réfraction. La loi de la conservation de l'impulsion est alors satisfaite même dans l'acte élémentaire.* Mais si on réalise une expérience ayant pour objet de déceler le photon (c'est ce qu'on appelle *mesure* en mécanique quantique), on décèlera soit un photon réfléchi, soit un photon transmis. Quoique toute mesure modifie l'état du système, la loi de la conservation de l'impulsion n'est pas violée. On ne peut prédire avec certitude l'état dans lequel on trouvera le photon — état de réflexion ou état de réfraction. On ne peut indiquer que les *probabilités* de chacun de ces états.

§ 6. Rayonnement de Vavilov-Cerenkov. Effet Doppler produit par une source lumineuse en mouvement dans un milieu

1. Soit une source de lumière en mouvement dans un milieu isotrope d'indice de réfraction n (ω). Les lois de la conservation de l'énergie et de l'impulsion pour une émission de la lumière se formulent exactement de la même façon que dans le cas du vide, i.e. sous la forme (4.7) où on conservera les mêmes notations. Il va de soi qu'on utilise un référentiel privilégié dans lequel le milieu est *immobile*. La mise en œuvre des seules lois de la conservation ne peut évidemment permettre de décider s'il se produira une émission de la lumière ou non. Pour répondre à cette question, il faut utiliser les équations électrody-

namiques plus détaillées. *Mais si les lois de la conservation ne sont pas respectées,* on peut affirmer qu'*aucune émission n'est possible.* Les lois de la conservation ne peuvent mettre en évidence le mécanisme du phénomène considéré; elles ne représentent que la condition *nécessaire* mais non suffisante de l'émission.

Nous traiterons les équations (4.7) exactement comme nous l'avons fait au § 4, à cette différence près que nous écrirons la relation entre l'énergie et l'impulsion sous la forme $p_{rad} = n\mathcal{E}_{rad}/c$. Cela donne

$$\mathcal{E}_0^2 - \mathcal{E}_0'^2 = \mathcal{E}_{rad}^2 (n^2 - 1) - 2\mathcal{E}\mathcal{E}_{rad} (\beta n \cos\vartheta - 1). \qquad (6.1)$$

Lorsque $n = 1$, cette relation se réduit, comme de juste, à la relation (4.8). Le cas qui nous concerne ici se distingue du cas du vide en ce que le phénomène fait intervenir, en plus de l'atome émetteur et du rayonnement, encore le *milieu* dans lequel se meut l'atome. On en tient compte de façon phénoménologique par l'intermédiaire de l'indice de réfraction de milieu.

2. Appliquons l'équation (6.1) au mouvement d'une particule dont l'*état interne n'est pas affecté par le mouvement*, ce qui signifie que la particule conserve constamment le même état quantique. S'il s'agit d'une particule stable, cette condition sera toujours satisfaite tant que l'énergie de son interaction avec le milieu sera insuffisante pour faire passer la particule dans un état de plus grande énergie. Il en est ainsi lorsqu'on envisage le mouvement d'un électron, d'un proton, ou même d'une particule élémentaire instable, à condition qu'elle ne se désintègre ou ne subit une transmutation quelconque pendant le trajet. Dans tous ces cas, on peut faire abstraction des états de liberté internes de la particule et la considérer comme dénuée de toute structure propre. Pendant le mouvement, la masse m_0 et l'énergie au repos $\mathcal{E} = m_0c^2$ de la particule restent constantes. En posant $\mathcal{E}_0' = \mathcal{E}_0$ dans (6.1), on trouve

$$\mathcal{E}_{rad}(n^2 - 1) - 2\mathcal{E}(\beta n \cos\vartheta - 1) = 0,$$

bien sûr, s'il y a émission ($\mathcal{E}_{rad} \neq 0$). Comme $\mathcal{E} = \mathcal{E}_0/\sqrt{1 - \beta^2}$, on obtient

$$\mathcal{E}_{rad} = \frac{2\mathcal{E}_0}{\sqrt{1-\beta^2}\,(n^2-1)} (\beta n \cos\vartheta - 1), \qquad (6.2)$$

et par suite

$$\cos\vartheta = \frac{1}{\beta n} + \frac{\sqrt{1-\beta^2}\,(n^2-1)}{2\beta n} \left(\frac{\mathcal{E}_{rad}}{\mathcal{E}_0}\right). \qquad (6.3)$$

Si la particule n'a émis qu'un seul quantum, $\mathcal{E}_{rad} = \hbar\omega$, et

$$\cos\vartheta = \frac{1}{\beta n(\omega)} + \frac{\sqrt{1-\beta^2}\,(n^2(\omega)-1)}{2\beta n(\omega)} \frac{\hbar\omega}{\mathcal{E}_0}. \qquad (6.4)$$

Le deuxième terme de (6.4) tient compte du recul qu'éprouve la particule lors de l'émission d'un quantum $\hbar\omega$. Si on néglige ce terme (on

peut le faire chaque fois qu'il est très petit en module devant l'unité), on retrouve la formule classique

$$\cos\vartheta = \frac{1}{\beta n(\omega)}. \qquad (6.5)$$

Il résulte de cette formule que $\cos\vartheta > 0$, ce qui signifie que la lumière ne peut être émise qu'*en avant*. Pour que la lumière émise ait une fréquence ω, il faut que $\beta n(\omega) > 1$. On obtient le même résultat avec la formule générale (6.4), puisque les deux termes du second membre y sont essentiellement positifs. Si on introduit la vitesse de phase de la lumière dans un milieu à l'aide de la formule $n(\omega) = c/v_{\text{phase}}$, on trouve

$$v > v_{\text{phase}}, \qquad (6.6)$$

ce qui signifie que *l'émission lumineuse ne peut avoir lieu que si la vitesse de la particule est supérieure à la vitesse de phase de la lumière dans le milieu.* Il s'ensuit qu'une particule se déplaçant dans le vide ne peut émettre de la lumière parce que dans ce cas $v_{\text{phase}} = c$. L'énergie de radiation provient de l'énergie cinétique de la particule, puisque lors de l'émission l'état interne de la particule ne change pas. La conséquence en est que le mouvement de la particule s'en trouve ralenti. Néanmoins, l'accélération elle-même n'exerce aucune influence sur le mécanisme du phénomène, ce qui permet de dire que l'émission lumineuse est possible lors du *mouvement uniforme* d'une particule chargée dans un milieu, à condition que la condition (6.6) soit satisfaite.

Le phénomène qui vient d'être décrit a été découvert en 1934 par P.A. Cerenkov (né en 1904) qui travaillait sous la direction de S.I. Vavilov (1891-1951). On l'appelle *effet Vavilov-Cerenkov*. La théorie classique de l'effet a été élaborée par I.E. Tamm (1895-1971) et I.M. Franck (né en 1908), et la théorie quantique fut l'œuvre de V.L. Ginsburg (né en 1916) (voir t. IV, § 38).

3. Considérons maintenant le cas où *l'état interne de la particule émettrice se modifie lors de son mouvement*. A titre d'exemple, on peut citer le cas d'un atome ou d'un ion. Dans ce cas, l'émission lumineuse se fait aux dépens de l'énergie interne et de l'énergie cinétique de l'atome. Soient ω la fréquence de la lumière émise par l'atome immobile et ω' la fréquence de la lumière émise par l'atome en mouvement (ce sont les mêmes notations que celles que nous avons utilisées au § 4, mais différentes de celles du sous-paragraphe précédent). D'autre part, écrivons

$$\mathcal{E}_0^2 - \mathcal{E}_0'^2 = (\mathcal{E}_0 + \mathcal{E}_0')(\mathcal{E}_0 - \mathcal{E}_0') = [2\mathcal{E}_0 - (\mathcal{E}_0 - \mathcal{E}_0')](\mathcal{E}_0 - \mathcal{E}_0').$$

En substituant cette expession dans (6.1), on négligera, pour simplifier, les carrés des petites quantités $\mathcal{E}_{\text{rad}}$ et $(\mathcal{E}_0 - \mathcal{E}_0')$. On verra plus loin que ces quantités sont, toutes deux, proportionnelles à h^2.

Ainsi, dans cette approximation, on tient compte des termes linéaires en $\hbar$ et on rejette les termes proportionnels à $\hbar^2$. De ce fait, la constante de Planck $\hbar$ s'élimine de l'expression définitive de la fréquence de la lumière émise et on retrouve pour l'effet Doppler le résultat classique. Aux petites corrections quantiques près, cette formule permet de tirer toutes les conclusions qui comptent. Après avoir effectué la substitution invoquée plus haut et ayant remarqué que $\mathscr{E} = \mathscr{E}_0/\sqrt{1-\beta^2}$, on obtient la formule classique

$$\mathscr{E}_{\text{rad}} = \frac{\sqrt{1-\beta^2}\,(\mathscr{E}_0 - \mathscr{E}_0')}{1-\beta n \cos\vartheta}. \tag{6.7}$$

Trois cas sont à envisager.

1) $1 - \beta n > 0$, i.e. la vitesse de la source de lumière est plus petite que la vitesse de phase de la lumière dans le milieu. On a donc *a fortiori* $1 - \beta n \cos\vartheta > 0$, ce qui signifie que le dénominateur de (6.7) est positif. Puisque lors de l'émission on a toujours $\mathscr{E}_{\text{rad}} > 0$, l'émission lumineuse n'est possible que si $\mathscr{E}_0 - \mathscr{E}_0' > 0$, ce qui exige qu'à la suite de l'émission l'atome passe de son niveau d'énergie initial sur un niveau d'énergie plus bas. Dans le cas envisagé, un atome non excité n'est pas susceptible d'émission. Pour passer de l'énergie à la fréquence, il suffit de remarquer que $\mathscr{E}_{\text{rad}} = \hbar\omega'$ et $\mathscr{E}_0 - \mathscr{E}_0' = \hbar\omega$, car pour un atome immobile, toute l'énergie interne correspondant à sa transition du niveau supérieur $\mathscr{E}_0$ sur le niveau inférieur $\mathscr{E}_0'$ est transformée en radiation. Dans ce cas, la formule (6.7) s'écrit sous la forme

$$\omega' = \frac{\omega\sqrt{1-\beta^2}}{1-\beta n\cos\vartheta} = \frac{\omega\sqrt{1-\beta^2}}{1-(v/v_{\text{phase}})\cos\vartheta}. \tag{6.8}$$

Elle ne diffère de la formule (4.10), correspondant à l'émission dans le vide, que par le remplacement de la vitesse c par la vitesse de phase de la lumière dans le milieu. Le milieu n'exerce aucune influence sur l'effet Doppler transversal ($\cos\vartheta = 0$). Cet effet n'est déterminé que par le ralentissement relativiste du temps (cf. t. IV, § 107).

2) $1 - \beta n < 0$, i.e. la vitesse de la source est plus grande que la vitesse de phase de la lumière dans le milieu. Dans ce cas, qui fut envisagé par V.L. Ginsburg et I.M. Franck, on parle de l'*effet Doppler superlumineux* ou *anomal*. On doit envisager ici deux possibilités.

Premièrement, la lumière est émise sous un angle ϑ tel que le dénominateur de (6.7) est positif, i.e. $\cos\vartheta < 1/(\beta n)$. Cela signifie que la radiation est émise dans une direction extérieure au *cône de Cerenkov*, i.e. au cône défini par l'équation $\cos\vartheta = 1/(\beta n)$. Comme dans le premier cas, l'émission n'est possible que si $\mathscr{E}_0 > \mathscr{E}_0'$, ce qui implique que l'atome doit passer d'un niveau d'énergie supérieur sur un niveau inférieur. La fréquence ω' est définie par la formule (6.8).

Deuxièmement, la lumière est émise sous un angle ϑ tel que le dénominateur de (6.7) est négatif, i.e. $\cos\vartheta > 1/(\beta n)$. Cela signifie que la radiation émise se propage à l'intérieur du cône de Cerenkov défini par $\cos\vartheta = 1/\beta n$. Pour que l'émission soit possible, il faut que $\mathscr{E}_0 < \mathscr{E}'_0$, ce qui implique le passage de l'atome d'un niveau inférieur sur un niveau d'énergie plus grande, i.e. son excitation. L'émission lumineuse et l'excitation de l'atome se font aux dépens de l'énergie cinétique de l'atome. Dans ce cas, $\mathscr{E}'_0 - \mathscr{E}_0 = \hbar\omega$, i.e. ω désigne la fréquence de la lumière qui serait émise par un atome immobile effectuant une transition du niveau supérieur $\mathscr{E}'_0$ sur le niveau inférieur $\mathscr{E}_0$. Cette fréquence est égale à la fréquence de la lumière que pourrait absorber un atome immobile pendant sa transition inverse du niveau inférieur $\mathscr{E}_0$ sur le niveau supérieur $\mathscr{E}'_0$. Quant à $\mathscr{E}_{rad}$, cette quantité est toujours définie par la relation $\mathscr{E}_{rad} = \hbar\omega'$. On arrive ainsi à la formule

$$\omega' = \frac{\omega\sqrt{1-\beta^2}}{\beta n\cos\vartheta - 1} = \frac{\omega\sqrt{1-\beta^2}}{(v/v_{\text{phase}})\cos\vartheta - 1}. \tag{6.9}$$

3) $1 - \beta n = 0$, i.e. la vitesse de l'atome est égale à la vitesse de phase de la lumière dans le milieu. Dans ce cas, si l'atome porte une charge, on voit apparaître le rayonnement de Vavilov-Cerenkov, que nous avons examiné dans le sous-paragraphe 2.

§ 7. Photons dans le champ de gravitation

1. En se plaçant à un point de vue quantique, on se propose d'examiner le changement de fréquence et l'incurvation d'un rayon lumineux, déterminés par le champ de gravitation. Nous avons déjà fait l'étude classique du changement de fréquence dans le t. I (§ 72) et dans le t. IV (§ 109) en nous fondant sur le principe de l'équivalence du champ de gravitation et du mouvement accéléré. Les résultats que nous y avons obtenus sont déduits ici de la loi de la conservation de l'énergie et de la relation entre l'énergie et la fréquence des photons: $\mathscr{E} = \hbar\omega$.

Selon la théorie de la relativité, toute énergie possède une masse, la masse d'inertie étant égale à la masse grave. Appliquons cette assertion à un faisceau lumineux limité, d'énergie $\mathscr{E}$, se propageant dans un champ de gravitation constant. Le potentiel gravitationnel du champ $\varphi(r)$ peut varier dans l'espace. Comme la lumière possède une masse grave $m = \mathscr{E}/c^2$, le champ de gravitation effectue du travail. Lorsque la lumière est transportée d'un point où le potentiel gravitationnel est φ en un point où le potentiel gravitationnel est $\varphi + d\varphi$, l'énergie de la lumière s'accroît de

$$d\mathscr{E} = -Gm\,d\varphi = -G\frac{\mathscr{E}}{c^2}\,d\varphi,$$

où G est la constante gravitationnelle. En intégrant cette équation entre les points *1* et *2*, on trouve

$$\ln \frac{\mathcal{E}_2}{\mathcal{E}_1} = \frac{G}{c^2} (\varphi_1 - \varphi_2). \tag{7.1}$$

C'est une relation générale qui n'inclut encore aucune hypothèse quantique. Elle est également valable en physique classique et en physique quantique. Mais pour en déduire une relation entre les fréquences, il faut faire appel à la relation entre l'énergie et la fréquence, qui est du ressort de la théorie quantique. Supposons que le faisceau lumineux ne se compose que d'un seul photon de fréquence ω. Dans ce cas, $\mathcal{E} = \hbar\omega$, et la relation (7.1) devient

$$\ln \frac{\omega_2}{\omega_1} = \frac{G}{c^2} (\varphi_1 - \varphi_2). \tag{7.2}$$

On voit que la constante de Planck ne figure plus dans le résultat final, et le résultat des calculs ne dépend pas de la valeur numérique de cette constante. Il doit en être ainsi chaque fois que le résultat final coïncide avec le résultat de la théorie classique.

Pour la première fois, la relation (7.2) a été mise à l'épreuve dans des observations astronomiques du déplacement des raies spectrales dans le champ de gravitation des étoiles. La découverte de l'effet Mössbauer (né en 1929) a incité Pound (né en 1919) et Rebka à vérifier cette relation dans les conditions terrestres (en 1960). Dans leurs expériences, ils mesuraient la variation de la fréquence de la lumière résultant d'un trajet vertical de 19,6 m dans le champ de pesanteur de la Terre. Cette question sera examinée en détail dans le volume consacré à la physique nucléaire.

2. On peut analyser l'incurvation des rayons lumineux dans un champ de gravitation en se fondant également sur le principe d'équivalence, sans aucun recours aux conceptions quantiques. C'est ainsi que cette question fut posée et résolue dès 1911 par A. Einstein avant que fut élaborée la théorie de la relativité générale. Notons cependant qu'*une solution complète de ce problème n'est possible que dans le cadre de la théorie de la relativité générale.* La solution qu'on en donne ci-après est exacte à un facteur numérique près. C'est une solution classique en fait, quoique présentant une apparence quantique.

Considérons un photon passant devant le Soleil ou un corps céleste massif de masse M. Si le champ de gravitation n'existait pas, la trajectoire du photon serait rectiligne. Le photon possède une masse inerte que nous dénoterons par m (il s'agit, bien sûr, de la masse relativiste, puisque la masse au repos du photon est nulle). En vertu du principe d'équivalence, la masse inerte est toujours égale à la masse grave. Par conséquent, le photon sera soumis à l'action de la force de gravitation $G \frac{Mm}{R^2}$, dirigée vers le centre du Soleil (R est la

distance au centre du Soleil). Nous avons déjà expliqué plus haut l'influence qu'exerce la composante tangentielle de cette force: elle détermine une variation de la fréquence de l'onde lumineuse. La composante normale de cette force provoque une incurvation du rayon lumineux. Il s'ensuit qu'un rayon lumineux passant près du Soleil doit s'incurver vers le centre de celui-ci.

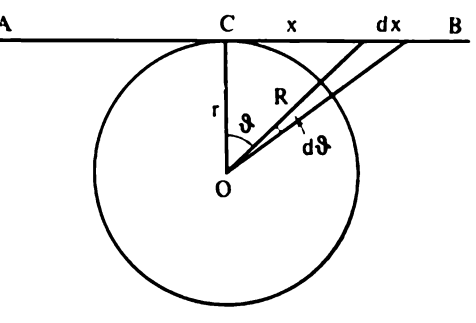


Fig. 13

Calculons l'angle de déviation du rayon lumineux. En l'absence de champ de gravitation, le rayon serait rectiligne (droite AB sur la fig. 13). Posons qu'en présence du champ de gravitation la forme du rayon sera peu différente de la droite AB. Le problème se ramène au calcul de l'impulsion $\int F_n dt$ de la force normale F_n appliquée au photon pendant toute la durée du mouvement. L'intégrale doit être calculée le long de la *trajectoire réelle* du photon, mais dans le cas considéré ici on appliquera la *méthode des perturbations* et on remplacera la trajectoire réelle par la trajectoire rectiligne AB non perturbée du photon. Supposons que le rayon non perturbé frôlerait le bord du Soleil; dans ce cas, conformément à la figure 13, on aurait

$$F_n = G\frac{Mm}{R^2}\cos\vartheta = G\frac{Mm}{r^2}\cos^3\vartheta,$$

$$x = r\,\mathrm{tg}\,\vartheta, \quad dx = \frac{r}{\cos^2\vartheta}d\vartheta,$$

où r est le rayon du Soleil. Par conséquent,

$$\int F_n\,dt = \int F_n\frac{dx}{c} = \frac{GMm}{cr}\int_{-\pi/2}^{+\pi/2}\cos\vartheta\,d\vartheta = \frac{2GMm}{cr}.$$

Cette impulsion de la force normale doit être égale à la variation de la quantité de mouvement du photon. Dans l'approximation envisagée, la quantité de mouvement du photon ne change qu'en direction, restant constante en module. Sa variation est égale à $mc\varphi$, où φ est l'angle dont tourne le rayon lumineux. En égalant ces deux expressions, on trouve

$$\varphi = 2\frac{GM}{c^2r}.$$

La théorie de la relativité générale fournit une valeur deux fois plus grande

$$\Phi = 4\frac{GM}{c^2 r}. \qquad (7.3)$$

Pour le Soleil $\Phi = 1{,}75''$, résultat conforme aux résultats expérimentaux.

§ 8. Expériences révélant les propriétés corpusculaires de la lumière

1. En 1925, A.F. Ioffé (1880-1960) et N.I. Dobronravov (1891-1949) ont confirmé par l'expérience la conception quantique de l'effet photoélectrique dont il a été question au § 2. Ils ont remplacé les rayons de lumière visible ou ultraviolette par des rayons X dont les quanta sont beaucoup plus grands. Un grain de poussière microscopique, qui acquérait des charges électriques lorsqu'on l'irradiait par des rayons X, était maintenu en équilibre dans le champ électrique d'un condensateur, analogue à celui qu'utilisait Millikan (1868-1953) pour déterminer la charge de l'électron. Tant que la charge portée par le grain de poussière était constante, celui-ci restait immobile, suspendu dans le champ électrique du condensateur, et pouvait être observé de côté à l'aide d'un microscope. Lorsqu'on l'irradiait avec un faible flux de rayons X, l'effet photoélectrique se manifestait. De temps en temps, un quantum X tombait sur le grain de poussière et libérait un électron. La charge du grain de poussière changeait de la valeur d'une charge électronique et l'équilibre s'en trouvait rompu. L'expérience montra que l'image quantique de l'effet photoélectrique correspondait bien au comportement du grain de poussière.

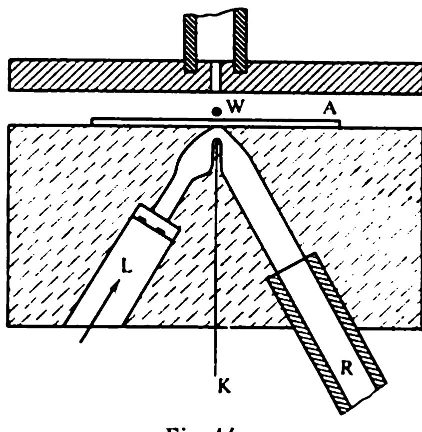


Fig. 14

Le schéma du dispositif expérimental de Ioffé et Dobronravov est représenté sur la figure 14. Une plaque d'ébonite épaisse comportait deux orifices L et R. A travers l'orifice R on évacuait l'air afin de rendre la cavité transparente aux rayons ultraviolets. A travers l'orifice L obturé par une fenêtre en quartz, on faisait pénétrer des rayons ultraviolets qui éclairaient l'extrémité supérieure K d'un fil d'aluminium de 0,2 mm de diamètre. Les photoélectrons ainsi produits étaient accélérés par une différence de potentiel de 12 000 V appliquée entre le fil d'aluminium et une feuille d'aluminium A

obturant la partie supérieure de la cavité. L'épaisseur de cette feuille d'aluminium ($\sim 5\cdot10^{-3}$ mm) était telle qu'elle n'absorbait pratiquement pas les rayons X qui y étaient excités par le freinage des électrons incidents. L'éclairement de la pointe du fil K devait être suffisamment faible pour que le nombre de photoélectrons et d'impulsions des rayons X qu'ils excitent soit d'environ 1000 par seconde. La feuille d'aluminium A faisait simultanément office d'armature inférieure du condensateur. A une distance de 0,02 cm de la feuille, un grain de bismuth W de $6\cdot10^{-5}$ cm de diamètre était maintenu en équilibre à l'état suspendu.

Les expériences ont montré qu'en moyenne toutes les 30 minutes l'équilibre du grain de bismuth était rompu, ce qui signifiait que c'était la fréquence moyenne de l'arrachement d'un électron par les rayons X. Pendant cet intervalle de temps, il se formait $N = 30 \times \times 60\cdot1000 = 1{,}8\cdot10^{6}$ impulsions de rayons X. Selon les conceptions classiques, l'énergie de chaque impulsion doit se propager dans toutes les directions sous forme d'une onde sphérique. Chacune des impulsions ne pourrait céder au grain de poussière qu'une partie infime de son énergie en raison de la petitesse de l'angle solide sous lequel le grain est vu de la région la plus proche de la feuille où sont excités les rayons X. En outre, cette portion infime d'énergie devrait être répartie parmi une multitude d'électrons du grain de bismuth. Dans ces conditions, il est absolument invraisemblable qu'en l'espace de 30 minutes une partie importante de l'énergie puisse être concentrée dans un seul électron afin qu'il puisse s'échapper. Il apparaît clairement que les résultats des expériences de Ioffé et Dobronravov ne sauraient être expliqués par la théorie ondulatoire classique. Par contre, ils se laissent parfaitement interpréter par la théorie quantique.

Un calcul élémentaire montre que l'interprétation quantique du phénomène est correcte d'un point de vue quantitatif. Si on pose que les directions des trajectoires d'un photon X émis sont équiprobables, la probabilité pour qu'un photon tombe sur le grain de bismuth soit égale à $\Omega/4\pi$, où Ω est l'angle solide sous lequel on voit le grain du point d'où sont émis les rayons X. Si la source émet par seconde, en moyenne, n photons, au bout d'un temps t elle aura émis $N = nt$ photons. De cette quantité $N\frac{\Omega}{4\pi} = nt\frac{\Omega}{4\pi}$ photons tomberont en moyenne sur le grain de bismuth. Le temps moyen t au cours duquel 1 photon atteint la cible se déduit de la condition $nt\frac{\Omega}{4\pi} = 1$, d'où

$$t = \frac{4\pi}{n\Omega}.$$

Dans les expériences de Ioffé et Dobronravov, $n = 1000\ \mathrm{s}^{-1}$, $\Omega = = \frac{\pi}{4}\left(\frac{6\cdot10^{-5}}{0{,}02}\right)^2$, en admettant que le grain est sphérique. En substi-

tuant ces valeurs dans la formule précédente, on trouve $t \approx 1800$ s= = 30 mn, valeur conforme aux résultats expérimentaux.

2. Selon les conceptions classiques, l'énergie radiante se propage simultanément dans toutes les directions. Mais il n'en sera pas ainsi si l'énergie se propage par quanta. Chaque quantum est émis indépendamment de tous les autres. Cet effet, s'il existe réellement, doit être d'autant plus net que les quanta sont grands. Bothe (1891-1957) utilisa (en 1924) les rayons X pour déceler cet effet.

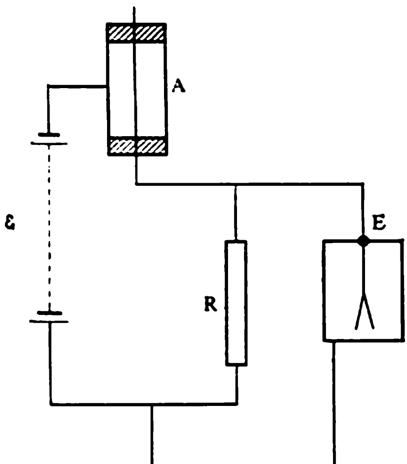


Fig. 15

Dans ses expériences, Bothe utilisait des compteurs de Geiger (1882-1945) pour déceler les quanta émis. Un compteur de Geiger se présente sous la forme d'un petit condensateur cylindrique A rempli de gaz à une pression de quelques mm Hg (fig. 15). L'armature intérieure du condensateur est un fil métallique fin tendu suivant l'axe du cylindre, ou bien une pointe métallique fixée sur une base isolante à l'intérieur du cylindre. Les armatures intérieure et extérieure du condensateur sont reliées l'une à l'autre à travers une pile B et une grande résistance R dont l'extrémité inférieure est mise à la terre. On maintient ainsi une différence de potentiel de 1000 V environ entre les armatures du condensateur. A proximité du fil ou de la pointe règne un champ électrique très inhomogène. Lorsqu'une particule quelconque (le photon y compris) pénètre dans le compteur, elle ionise les molécules du gaz. Les électrons et les ions qui se sont formés sont accélérés dans ce champ et ionisent par choc les molécules du gaz. Les nouveaux ions et électrons résultant de cette ionisation sont accélérés par le champ, et le processus d'ionisation se poursuit. Il se forme une avalanche d'électrons et d'ions et le compteur est traversé par une impulsion de courant électrique d'intensité notable. La répartition de la différence de potentiel initialement appliquée s'en trouve modifiée, la majeure partie sera appliquée à la résistance R, tandis que la tension appliquée au condensateur du compteur tombe presque à zéro. A ce moment, le courant traversant le compteur tombe à zéro, la tension appliquée aux bornes du condensateur augmente et le compteur revient à son régime de travail. On peut déceler l'impulsion de courant qui a traversé le compteur à l'aide d'un électromètre E sensible, relié à l'armature intérieure du condensateur du compteur. En détectant les impulsions on peut en-

registrer (compter) chaque particule qui provoque l'ionisation primaire du gaz. Les compteurs de Geiger modernes sont munis de dispositifs d'enregistrement automatiques. Les compteurs de Geiger ont joué un rôle important et sont toujours utilisés en physique nucléaire, en physique des rayons cosmiques, en physique des particules élémentaires, etc.

Dans les expériences de Bothe, on utilisait deux compteurs C_1 et C_2 à action rapide (fig. 16), qui permettaient d'enregistrer chaque quantum X qui les traversait. Entre les compteurs on disposait, en position symétrique, une mince feuille A en fer ou en cuivre. Cette feuille était irradiée de côté par des rayons X durs R et devenait elle-même source d'un rayonnement X caractéristique (fluorescence X).

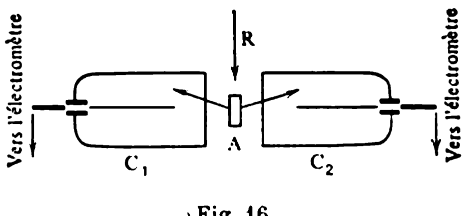


Fig. 16

L'arrivée d'un quantum X dans l'un des compteurs était presque instantanément (en moins de 0,001 s) marquée par un tressaillement du fil de l'électromètre. Ces tressaillements étaient enregistrés automatiquement sur une bande de papier mobile. Si, conformément à la théorie classique, la source A émettait des ondes dans toutes les directions, les compteurs seraient déclenchés simultanément et en concordance (avec de petites variations aléatoires). D'autre part, si la source A émettait des quanta X à gauche et à droite, de façon désordonnée et indépendante, le déclenchement des compteurs serait également désordonné et indépendant. Les expériences ont montré que tout se passe comme si le rayonnement X de fluorescence se propageait sous forme de quanta dirigés aléatoirement vers l'un ou l'autre compteur. Ce comportement correspond bien aux conceptions quantiques.

3. Dans les deux études expérimentales que nous venons de décrire, il s'agissait en fait de *phénomènes de fluctuation* se produisant sous l'action de faibles flux de rayons X. Il est fort difficile d'observer des phénomènes analogues avec la lumière visible, pour la raison que les quanta de lumière sont très petits dans cette région du spectre. La sensibilité des méthodes d'enregistrement objectives est encore insuffisante pour pouvoir enregistrer des quanta isolés de lumière visible. De faibles flux de lumière furent observés par S. I. Vavilov. Il mit à profit le fait que les régions périphériques de la rétine de l'œil humain (les bâtonnets) acquièrent une sensibilité suffisante à une lumière faible après un séjour prolongé dans l'obscurité. Vavilov profita également de l'existence d'un seuil de perception visuelle. Lorsque le nombre de quanta, parvenant lors d'une brève impulsion lumineuse jusqu'aux régions périphériques de la rétine, est supérieur à une valeur minimale n_0, l'œil enregistre l'impulsion

lumineuse, et ne la décèle pas si ce nombre est inférieur à n_0. Selon les estimations de Vavilov, n_0 est de l'ordre de quelques dizaines ou même de quelques quanta seulement. L'existence du seuil de perception visuelle a été systématiquement utilisée par Vavilov dans ses études de la luminescence, et avec Cerenkov, dans l'étude de l'effet Vavilov-Cerenkov. Mettant à profit ces propriétés de l'œil, Vavilov mit au point une méthode sensible d'observations visuelles des fluctuations d'intensité de la lumière visible, dans le cas de faibles flux lumineux. Après des années d'étude de ces fluctuations, il arriva à la conclusion qu'elles se produisent de façon analogue aux fluctuations du nombre de molécules de gaz, i.e. comme si la lumière était composée de portions finies ou quanta de lumière.

CHAPITRE II

STRUCTURE, NIVEAUX D'ÉNERGIE ET SPECTRES DES ATOMES

§ 9. Modèle nucléaire de l'atome et expériences de Rutherford

1. Dans le premier quart du XXᵉ siècle, on arriva à établir que tout atome se compose d'un *noyau* portant une charge positive et d'une *enveloppe électronique*. Les dimensions linéaires du noyau sont de l'ordre de 10^{-13} à 10^{-12} cm, et celles de l'atome, déterminées par l'enveloppe électronique, sont 10^5 fois plus grandes. Néanmoins, la presque totalité de la masse de l'atome (plus de 99,95 %) est concentrée dans le noyau. Cela tient à ce que le noyau se compose de protons et de neutrons qui sont des particules lourdes, tandis que l'enveloppe électronique se compose uniquement d'électrons qui sont « légers » ($m_p = 1836{,}15 m_e$, $m_n = 1838{,}68 m_e$). Le nombre d'électrons constituant le cortège électronique d'un atome neutre est égal à la charge du noyau, à condition de prendre pour unité de charge la charge élémentaire (valeur absolue de la charge élecrique de l'électron). L'enveloppe électronique de l'atome peut perdre ou acquérir des électrons; dans ce cas, l'atome devient porteur d'une charge électrique et se transforme en un *ion* positif ou négatif.

Les propriétés chimiques des atomes sont déterminées par l'enveloppe électronique, plus exactement par ses électrons périphériques. Ces derniers sont plus faiblement liés à l'atome et de ce fait sont plus sensibles aux actions électriques exercées par les électrons périphériques des atomes voisins. La même chose vaut pour les forces d'attraction et de répulsion s'exerçant entre les atomes et les molécules neutres (forces moléculaires). D'autre part, les neutrons et les protons sont fortement liés dans le noyau. Pour pouvoir exercer une action sur le noyau, on doit mettre en œuvre des forces des millions de fois plus grandes que celles qui suffisent pour arracher les électrons périphériques de l'atome. Néanmoins, la structure et les propriétés de l'enveloppe électronique dépendent, en fin de compte, du champ électrique créé par le noyau de l'atome.

2. En admettant que ce modèle de l'atome correspond à la réalité, les atomes devraient se laisser traverser aisément par des particules. Lenard le montra en utilisant un faisceau d'électrons, mais c'est Rutherford (1871-1937) qui, en 1911, démontra par l'expérience que ce modèle de l'atome était correct. C'est pour cela qu'on l'appelle *modèle d'atome de Rutherford*. Sur suggestion et sous la direction de Rutherford, deux de ses disciples, Geiger et Marsden (1889-1970),

étudièrent quantitativement la diffusion des particules α émises par des substances radioactives. Un faisceau parallèle de particules α se propageant dans le vide tombait sur une feuille métallique de faible épaisseur et subissait une diffusion. La détection des particules α diffusées était visuelle : au point d'impact, sur un écran fluorescent de sulfure de zinc, chaque particule α produisait une scintillation. A l'aide d'une loupe ou d'un microscope on observait dans l'obscurité les scintillations produites et on les dénombrait.

On constata que la grande majorité des particules α étaient diffusées dans un angle de 1 à 3°. La distribution angulaire de ces particules était décrite de façon satisfaisante par la courbe des erreurs de Gauss (1777-1855). Mais on décela aussi des particules α dont l'angle de déviation était beaucoup plus grand, atteignant parfois 150°. Le nombre relatif de particules fortement déviées était très petit. Par exemple, en dirigeant un faisceau de particules α issues de RaC *) sur une feuille de platine, sur 8000 particules on ne trouvait en moyenne qu'une seule qui était déviée d'un angle supérieur à 90°. Si les déviations sous grands angles résultaient d'une accumulation de déviations dues au hasard, ce rapport devrait être encore plus petit.

Rutherford arriva à la conclusion que chaque grande déflexion résultait d'un *acte d'ineraction élémentaire* entre un centre de force pratiquement ponctuel et une particule α passant à proximité. Ce centre de force fut identifié avec le *noyau* positivement chargé d'un atome. La particule α elle-même est également un noyau atomique — le noyau de l'atome d'hélium. Rutherford démontra en effet qu'on peut obtenir une particule α en ionisant deux fois de suite un atome d'hélium. L'interaction électrostatique de ces deux noyaux provoque la diffusion sous grands angles des particules α.

Ces considérations sont confirmées par les photographies des traces que laissent les particules α dans la chambre de Wilson. Généralement, l'extrémité de la trace d'une particule α ne présente aucune particularité, mais on observe parfois des traces présentant en fin de course des brisures ou des bifurcations. Un exemple en est donné sur la figure 17, qui représente le choc d'une particule α avec un noyau. Après le choc, la direction du mouvement de la particule α s'est notablement modifiée, tandis que le noyau mis en mouvement a laissé une nouvelle trace ; celle-ci forme avec la trace de la particule α une « fourche ».

3. Rutherford élabora une théorie quantitative de la diffusion des particules α. Dans cette théorie, il appliqua la *loi de Coulomb* (1736-1806) à l'interaction de la particule α avec le noyau, quoique la validité de cette loi n'ait pas été vérifiée par l'expérience pour une distance de séparation de l'ordre de 10^{-12} cm, distance à laquelle la

---

*) En réalité, les particules α étaient émises par RAC', qui est un produit à courte vie de RaC (la période de demi-vie de RaC' est de 10^{-6} s).

particule α peut passer près du noyau (voir problème n° 1 à la fin de ce paragraphe). Bien entendu, Rutherford traitait le mouvement de la particule α dans le champ du noyau, conformément à la mécanique *classique*. On peut justifier par la mécanique quantique (qui s'est constituée peu de temps après, cf. § 20) l'application de la mécanique classique au problème considéré par Rutherford. Enfin, il posait que la masse du noyau était assez grande par rapport à celle

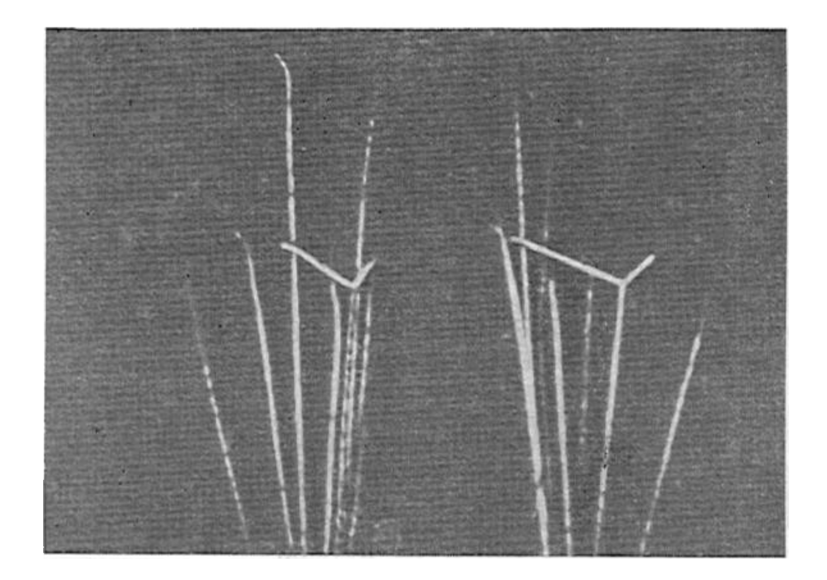


Fig. 17

de la particule α pour admettre que le noyau était immobile. On se passe aisément de cette dernière hypothèse en remplaçant la masse de la particule α par la *masse réduite*.

Rutherford utilisait dans ces expériences des feuilles métalliques très fines, d'une épaisseur de 10^{-5} à 10^{-4} cm, afin de ne pas avoir à tenir compte de l'éventualité de chocs multiples des particules α avec les noyaux atomiques dans l'interprétation des grands angles de diffusion. En effet, la probabilité de ce qu'une grande déviation résulterait de deux, et à fortiori, de plusieurs chocs successifs est absolument négligeable. La probabilité de diffusion sous grands angles, due aux électrons, est également négligeable en raison de la faible masse de ces derniers. Les chocs multiples avec les noyaux et les électrons gravitant autour de ceux-ci ne jouent un rôle que lorsque l'angle de diffusion est très petit. Nous ne tiendrons pas compte des très petits angles de diffusion. En ne retenant que l'interaction d'une

particule α avec un seul noyau, celui qui se trouve le plus près du projectile, on arrive au *problème de deux corps*. La particule α considérée traversant la feuille se trouve en effet à des distances beaucoup plus grandes de tous les autres noyaux, ce qui permet de négliger leurs interactions avec la particule α. La théorie de Rutherford s'applique aux grandes déviations lorsque chacune d'elles est due au champ électrique d'un *seul noyau*; en comparaison de cette déviation, toutes

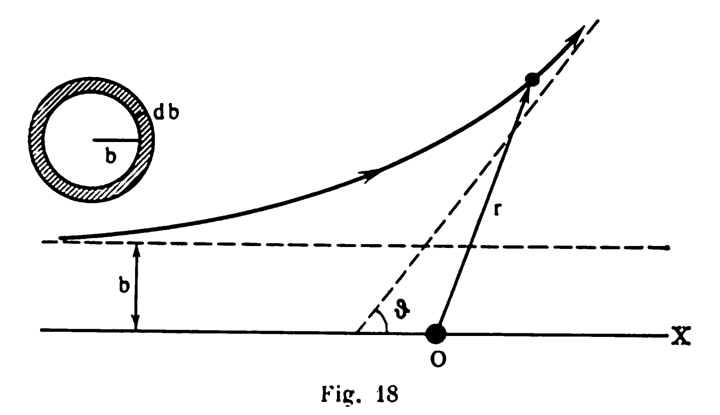


Fig. 18

les autres, prises ensemble, sont parfaitement négligeables. La diffusion ainsi définie est dite *diffusion de Rutherford*. C'est une diffusion élastique en ce sens que l'énergie cinétique de la particule α ne change pas du fait de la diffusion, i.e. n'est pas dissipée pour exciter les atomes et à fortiori pour exciter les noyaux.

Ainsi formulé, le problème est analogue au *problème de Kepler* (1571-1630) concernant le mouvement des planètes autour du Soleil. Dans les deux cas, la force d'interaction s'exerçant entre les corps est une force centrale qui varie en raison inverse du carré de la distance entre les corps. Dans le cas des planètes, il s'agit d'une force d'attraction, et dans celui de la particule α, il s'agit d'une force de répulsion. Cette différence se traduit par ce qu'une planète peut décrire, suivant la valeur de son énergie totale, une trajectoire elliptique ou hyperbolique, tandis que la particule α ne peut se déplacer que suivant une hyperbole. Cette différence n'affecte absolument pas le calcul mathématique. L'angle de diffusion ϑ de la particule α est égal à l'angle que forment les asymptotes de sa trajectoire hyperbolique (fig. 18). La valeur de cet angle est donnée par la formule (t. I, § 58)

$$\operatorname{cotg}\frac{\vartheta}{2}=\frac{mbv^2}{2Ze^2}, \tag{9.1}$$

où m est la masse de la particule α, v sa vitesse à l'« infini »; i.e. loin du noyau, Ze la charge du noyau, $2e$ la charge de la particule

α, égale au double de la charge élémentaire e. (Le nombre Z est couramment appelé *charge du noyau*, entendant par là qu'on prend pour l'unité la charge élémentaire e.) b désigne le *paramètre de choc*, i.e. la longueur de la perpendiculaire abaissée du noyau sur la trajectoire rectiligne non perturbée de la particule α (ou, ce qui revient au même, abaissée sur la tangente à la trajectoire réelle de la particule α lorsque celle-ci se trouvait à distance infinie du noyau).

4. Dans le domaine des phénomènes atomiques, ce n'est pas la formule (9.1) qui est susceptible de vérification expérimentale, mais les *résultats statistiques* auxquels elle conduit. Introduisons la notion de *section efficace différentielle de diffusion*. Notons I l'intensité d'un faisceau plan-parallèle de particules α tombant sur le noyau, i.e. le nombre de particules α du faisceau passant dans l'unité de temps à travers une aire unité normale au faisceau. Le nombre de particules α traversant une surface élémentaire $d\sigma$, normale au flux, est $dN_1 = d\sigma I$. Après diffusion, ces particules sont contenues dans un angle solide élémentaire $d\Omega$. Il va de soi que la valeur de l'angle solide $d\Omega$ et la direction de son axe dépendent de la position et de l'aire de la surface élémentaire $d\sigma$. Il s'ensuit que dN_1 représente aussi le nombre de particules α diffusées par le noyau dans l'unité de temps dans l'angle solide $d\Omega$. Le rapport de dN_1 à I est égal à $d\sigma$ et a la dimension d'une surface; c'est ce rapport qu'on appelle *section efficace différentielle d'un noyau provoquant la diffusion des particules α dans l'angle solide $d\Omega$*. Cette notion est utilisée non seulement dans l'étude de la diffusion des particules α, mais dans celle de *n'importe quelles particules*, ainsi que d'*autres processus* concernant les particules. Ainsi, par définition

$$d\sigma = \frac{dN_1}{I}, \tag{9.2}$$

ce qui signifie que *la section efficace différentielle de diffusion est le rapport du nombre de particules diffusées par l'atome dans l'unité de temps dans l'angle solide $d\Omega$ à l'intensité I des particules incidentes.*

Calculons la section différentielle de diffusion des particules α par un noyau atomique. Le problème se ramène au calcul de l'aire $d\sigma$ de l'élément de surface, en passant par lequel la particule α pénètre après diffusion dans un angle solide $d\Omega$ donné. Faisons coïncider l'axe X avec la trajectoire rectiligne de la particule α dont le paramètre de choc $b = 0$ (le choc d'une telle particule avec un noyau serait frontal). En adoptant une symétrie cylindrique, on remplace, pour simplifier, $d\sigma$ par l'aire d'une couronne $d\sigma = 2\pi b\, db$, normale à la direction du flux de particules. Le rayon intérieur de cette couronne est b, son rayon extérieur est $b + db$ et son centre se trouve sur l'axe X (fig. 18, en haut à gauche). A l'intervalle b, $b + db$ correspond un intervalle d'angles de diffusion ϑ, $\vartheta + d\vartheta$, db étant donné

selon la formule (9.1) par

$$db = \frac{Ze^2}{mv^2} \frac{d\vartheta}{\sin^2(\vartheta/2)}.$$

En introduisant l'angle solide $d\Omega = 2\pi \sin \vartheta \, d\vartheta$, dans lequel sont diffusées les particules α qui ont traversé la couronne, il est aisé d'obtenir

$$d\sigma = \left(\frac{Ze^2}{mv^2}\right)^2 \frac{d\Omega}{\sin^4(\vartheta/2)}. \tag{9.3}$$

Sous cette forme, la formule (9.3) s'applique à *n'importe quelle* surface élémentaire $d\sigma$. C'est la *formule de Rutherford.*

5. Avant de poursuivre notre exposé, il convient d'introduire la notion de *section efficace totale de diffusion* ou de tout autre processus. On la définit comme le *rapport du nombre total de particules, ayant participé au processus dans l'unité de temps, à l'intensité du faisceau de particules incidentes*. On peut déduire la section totale σ par intégration de la section différentielle $d\sigma$ sur toutes les valeurs possibles de $d\Omega$. Lorsqu'il s'agit de la diffusion de particules α, on doit poser dans la formule (9.3) $d\Omega = 2\pi \sin \vartheta \, d\vartheta$ et intégrer entre $\vartheta = 0$ et $\vartheta = \pi$ (voir problème n° 2 à la fin du paragraphe). On trouve ainsi $\sigma = \infty$. On conçoit que ce résultat est normal en remarquant que plus l'élément de surface $d\sigma$ est éloigné de l'axe X, plus l'angle de diffusion ϑ est petit. Les particules qui traversent des éléments de surface lointains ne sont pratiquement pas déviées et leurs trajectoires sont voisines de l'angle de diffusion $\vartheta = 0$. L'aire totale de tous ces éléments de surface et, par voie de conséquence, le nombre total de particules diffusées sont infiniment grands. Il s'ensuit que la section totale de diffusion l'est également. Ce résultat ne présente qu'un intérêt tout formel, puisque la formule de Rutherford (9.3) ne s'applique pas aux petits angles de diffusion.

6. Cherchons à mettre la formule (9.3) sous une forme susceptible de vérification expérimentale. Les actes de diffusion des particules α par des atomes différents sont *indépendants* les uns des autres. On en conclut que si n est le nombre de noyaux (d'atomes) contenus dans l'unité de volume, le nombre de particules α pouvant être diffusées par un volume V dans l'unité de temps dans un angle solide $d\Omega$, est donné par l'expression

$$dN = VnI \, d\sigma = VnI \left(\frac{Ze^2}{mv^2}\right)^2 \frac{d\Omega}{\sin^4(\vartheta/2)}. \tag{9.4}$$

Ecrite sous la forme (9.4), la formule de Rutherford a pu être justifiée par l'expérience. Les résultats expérimentaux ont montré notamment que pour $d\Omega$ constant, la quantité $dN \sin^4(\vartheta/2)$ est constante, i.e. ne dépend pas de l'angle de diffusion ϑ, comme l'impose la formule (9.4).

La justification expérimentale de la formule de Rutherford peut être considérée comme une *preuve indirecte de la validité de la loi de Coulomb* à des distances aussi petites que celles auxquelles peuvent se trouver les centres des particules α et des noyaux avec lesquels elles entrent en interaction. Une autre preuve de la validité de cette loi aux courtes distances est donnée par les résultats des expériences de Blackett (1897-1974) concernant la diffusion des particules α par les gaz. Dans ces expériences, on photographiait les traces des trajectoires des particules α dans une chambre de Wilson, on mesurait leurs déviations angulaires et on calculait la fréquence d'apparition des différents angles de diffusion. Ces expériences apportèrent une nouvelle confirmation de la justesse de la formule de Rutherford, quoique leur principal objectif était la vérification de la loi de Coulomb. On s'assura qu'à des distances entre les centres des particules α et les noyaux, égales à $3\cdot10^{-12}$-$5\cdot10^{-10}$ cm pour l'air et à $7\cdot10^{-12}$-10^{-9} cm pour l'argon, la loi de Coulomb se trouvait vérifiée par l'expérience. On ne peut cependant en inférer que cette loi est valable si petites que soient les distances entre les centres des noyaux en interaction. Des expériences sur la diffusion élastique de noyaux légers. accélérés dans des accélérateurs de particules, par des noyaux légers fixes, ont montré que lorsque la distance indiquée devenait égale ou inférieure à 10^{-12} cm, on observait des écarts à la loi de Coulomb. A ces distances commencent à se manifester les *forces d'attraction nucléaires* qui masquent les forces coulombiennes de répulsion entre les noyaux.

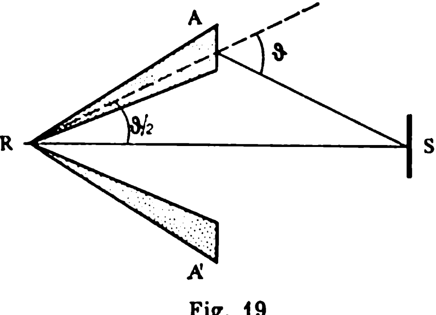


Fig. 19

7. On peut utiliser la formule (9.4) pour déterminer la charge du noyau. Pour cela il faut mesurer dN et I, après quoi on peut calculer Z. puisque toutes les autres quantités figurant dans la formule (9.4) peuvent être considérées connues. La principale difficulté de cette méthode consiste en ce que les quantités dN et I sont très différentes. Dans les premières expériences, on les mesurait à l'aide d'appareils de mesure différents, autrement dit dans des conditions différentes, ce qui introduisait des erreurs importantes. Dans les expériences de Chadwick (1891-1974), ce défaut de la méthode fut éliminé en procédant de la façon suivante. La feuille diffusante AA' avait la forme d'une couronne (fig. 19) ; le corps radioactif R (source de particules α) et l'écran fluorescent S en ZnS étaient disposés sur l'axe de la couronne à égales distances de cette dernière. Pour pouvoir dénombrer les scintillations produites par les particules α

diffusées par la feuille métallique, on obturait l'ouverture de la couronne AA' par un écran opaque aux particules α. Pour la mesure de I, on dénombrait les scintillations produites lorsque la couronne métallique était recouverte et son ouverture était libre. Comme dans ce dernier cas le nombre de scintillations était très grand, on le réduisait en plaçant devant l'écran S un disque tournant comportant une fente étroite. Connaissant la largeur de la fente et ayant dénombré le nombre de scintillations, on peut calculer I. Chadwick trouva pour le platine $Z = 77{,}4$, pour l'argent $Z = 46{,}3$, pour le cuivre $Z = 29{,}3$. Les numéros atomiques de ces éléments dans le tableau périodique de Mendéléev (1834-1907) sont respectivement égaux à 78, 47, 29. Ces expériences confirmèrent le résultat connu de Moseley (1887-1915) selon lequel la charge Z du noyau coïncide avec le numéro atomique de l'élément (cf. § 48).

8. Revenons au modèle de l'atome que justifièrent les expériences de Rutherford et demandons-nous si le noyau atomique et son cortège électronique peuvent former un système stable, ce qui est certainement vrai pour les atomes. Même si c'était possible, toutes ces particules ne pourraient rester au repos, car dans ce cas on aurait affaire à un système électrostatique de charges pratiquement ponctuelles entre lesquelles s'exerceraient des forces coulombiennes; or, selon le théorème de Earenshaw (cf. t. III, § 9), un tel système est instable. Les forces coulombiennes varient en raison inverse du carré de la distance entre les particules; les forces gravitationnelles entre les corps d'un système planétaire varient suivant la même loi. Or, la stabilité du système planétaire est assurée par la rotation des planètes autour du Soleil. C'est pour cela que Rutherford arriva tout naturellement au *modèle planétaire de l'atome* où les électrons tournent autour du noyau.

Or, selon l'électrodynamique classique, le déplacement de toute charge électrique fait varier le champ électromagnétique créé par cette charge. *Une charge électrique animée d'un mouvement accéléré émet des ondes électromagnétiques.* Un électron tournant possède une accélération et doit donc émettre constamment des ondes électromagnétiques et perdre de l'énergie. Perdant de l'énergie, l'électron devrait se rapprocher de plus en plus du noyau et finalement tomber sur celui-ci (voir problème n° 4 à la fin du paragraphe). Ainsi, même en admettant que les électrons sont en mouvement, on aboutit à un modèle d'atome qui est *instable*. On pourrait supposer que la loi de Coulomb et les lois d'électrodynamique qui définissent le champ électromagnétique seraient en défaut dans le cas des particules élémentaires et des faibles distances mises en jeu. On pourrait tenir compte des forces nucléaires et introduire des forces hypothétiques, encore inconnues, assurant la stabilité de l'atome. Mais le problème n'en serait pas résolu pour autant. Conformément aux principes de la mécanique classique, quelles que soient les forces que l'on fait

intervenir, le spectre d'émission de l'atome devrait comporter plusieurs fréquences fondamentales et leurs harmoniques. Or les données expérimentales s'expriment par le *principe de combinaison de Ritz* (1878-1909) (cf. § 11). On est amené à constater que la mécanique et l'électrodynamique classiques ne peuvent expliquer l'*existence des atomes* en tant que systèmes stables composés de noyaux atomiques et d'électrons. Ce problème n'a été résolu qu'à l'aide de la mécanique quantique.

PROBLÈMES

1. A quelle distance r peuvent se rapprocher, lors d'un choc frontal, les centres d'une particule α ayant une énergie $\mathscr{E} = 6$ MeV et d'un atome d'or immobile? La charge du noyau de l'atome d'or est $Z = 79$.

Réponse. $r = 2Ze^2/\mathscr{E} = 3{,}8\cdot 10^{-12}$ cm.

2. Quelle doit être la section $\Delta\sigma$ de la diffusion de Rutherford d'une particule α par un noyau atomique pour que l'angle de diffusion soit supérieur à ϑ?

Réponse.

$$\Delta\sigma = 4\pi \left(\frac{Ze^2}{mv^2}\right)^2 \operatorname{cotg}^2 \frac{\vartheta}{2}. \tag{9.5}$$

3. On a utilisé dans des expériences de diffusion des particules α une feuille de platine d'épaisseur $\delta = 8\cdot 10^{-5}$ cm. Le parcours dans l'air des particules α émises par du radium C' a été trouvé égal à 6,96 cm (à 15 °C et 760 mm Hg). En traçant le graphique représentant la relation entre la longueur du parcours et l'énergie cinétique $\mathscr{E}$ des particules α, on trouve $\mathscr{E} \approx 5{,}9$ MeV. Calculer la quantité relative de particules α dont l'angle de diffusion est supérieur à $\vartheta =$ $= 90°$. Densité du platine $\delta = 21{,}5$ g/cm³, charge du noyau $Z = 78$, masse atomique $A = 195$.

Solution. D'après la formule (9.5), la section de diffusion pour les angles $\vartheta > \pi/2$ est

$$\Delta\sigma = 2\pi \left(\frac{Ze^2}{\mathscr{E}}\right)^2,$$

vu que l'énergie cinétique d'une particule α est $\mathscr{E} = {}^1/_2 mv^2$. Le nombre de noyaux diffusants contenus dans la feuille est $N = S\delta n$, S étant l'aire totale de la feuille et n le nombre de noyaux contenus dans l'unité de volume. L'aire revenant à un noyau est $s = S/N = 1/(\delta n)$. Dans les angles $\vartheta > \pi/2$ ne parviennent que les particules α qui n'ont subi chacune qu'un seul choc avec les noyaux. Ces chocs uniques sont *indépendants* les uns des autres; par suite la quantité relative de particules diffusées sous un angle supérieur à $\pi/2$ est

$$\frac{\Delta\sigma}{s} = 2\pi \left(\frac{Ze^2}{\mathscr{E}}\right)^2 \delta n.$$

En y substituant $n = \frac{\rho}{A} N_{Av}$, N_{Av} étant le nombre d'Avogadro, on obtient

$$\frac{\Delta\sigma}{s} = 2\pi \left(\frac{Ze^2}{\mathscr{E}}\right)^2 \frac{\rho}{A}\,\delta N_{Av} = \frac{1}{8200}.$$

4. Au bout de quel intervalle de temps t un électron tournant autour d'un proton, sur une orbite circulaire de rayon $a_0 = 0{,}53\cdot 10^{-8}$ cm, tomberait sur le noyau par suite de pertes d'énergie par rayonnement, si on utilise pour le calcul les lois de la mécanique et de l'électrodynamique classiques?

Solution. Tout en tenant compte de l'émission d'un rayonnement, on peut utiliser, pour un calcul approché de l'accélération de l'électron, la formule

$|\dot{v}| = v^2/a$, où a est la distance variable entre l'électron et le noyau. Dans le cas d'un mouvement circulaire, $mv^2/a = e^2/a^2$, d'où

$$\frac{mv^2}{2} = \frac{e^2}{2a}.$$

L'énergie totale de l'électron est

$$\mathscr{E} = \frac{mv^2}{2} - \frac{e^2}{a} = -\frac{e^2}{2a}.$$

La perte d'énergie pour la radiation par unité de temps est

$$-\frac{d\mathscr{E}}{dt} = \frac{2}{3}\frac{e^2}{c^3}\dot{v}^2 = \frac{2}{3}\frac{e^2}{c^3}\frac{v^4}{a^2} = \frac{2}{3}\frac{e^2}{m^2c^3a^4}.$$

En y substituant l'expression de $\mathscr{E}$, on trouve

$$\frac{d}{dt}\left(\frac{1}{a}\right) = \frac{4}{3}\frac{e^4}{m^2c^3a^4},$$

d'où

$$a_0^3 - a^3 = 4\frac{e^4}{m^2c^3}t.$$

En posant $a = 0$, on trouve le temps de chute cherché

$$t = \frac{m^2c^3a_0^3}{4e^4} = 1{,}6\cdot 10^{-11} \text{ s}.$$

§ 10. Détermination de la charge du noyau à partir de données sur la diffusion des rayons X

1. Au § 3, en analysant la diffusion des rayons X par les électrons libres, on n'a prêté attention qu'à la variation de la longueur d'onde en fonction de l'angle de diffusion. Pour traiter ce problème, il a suffi d'avoir recours à une théorie quantique simple, fondée sur les seules lois de la conservation de l'énergie et de l'impulsion. Mais lorsqu'il s'agit de déterminer l'intensité et la polarisation de la radiation diffusée suivant différentes directions, il est nécessaire de faire appel au système complet d'équations de l'électrodynamique quantique et de la mécanique quantique relativiste. Ce problème fut résolu, en 1929, par O. Klein (1894-1977) et Nishina (1890-1951), puis, d'une façon plus rigoureuse, par I. E. Tamm en 1930. L'étude de cette question sort largement du cadre de notre cours, mais la théorie classique simple permet d'arriver à un résultat correct dans le cas limite où l'énergie du quantum incident $h\nu$ est petite devant l'énergie propre de l'électron m_ec^2 (on peut dire aussi : lorsque la longueur d'onde λ est grande devant la longueur de Compton λ_C). Ce cas présente un intérêt réel puisque son étude permet d'élaborer, pour les éléments légers, une méthode indépendante de détermination de la charge Z des noyaux. Effectuons ces calculs pour le cas où $h\nu \ll m_ec^2$, tout en posant que les quanta X sont assez durs pour que

leur énergie soit grande devant l'énergie de liaison des électrons ; on pourra donc poser que les électrons sont libres. On ne peut satisfaire à ces deux conditions que dans le cas d'*éléments légers*.

2. Dans un champ électrique monochromatique $E = E_0 \cos \omega t$, un électron libre prend une accélération $\ddot{x} = -(e/m)E = -(e/m)E_0 \cos \omega t$ (à condition de négliger l'action du champ magnétique). Selon l'électrodynamique classique, un électron accéléré émet une radiation et perd son énergie sous forme de radiation diffusée. L'énergie dissipée dans l'unité de temps par l'électron est donnée par

$$-\frac{d\mathscr{E}}{dt} = \frac{2}{3}\frac{e^2}{c^3}\ddot{x}^2 = \frac{2}{3}\frac{e^4}{m^2c^3}E_0^2\cos^2\omega t$$

(voir t. III, § 141). La moyenne temporelle de cette quantité

$$-\left(\overline{\frac{d\mathscr{E}}{dt}}\right) = \frac{1}{3}\frac{e^4}{m^2c^3}E_0^2.$$

On peut négliger la diffusion par les noyaux atomiques lourds puisqu'on verra apparaître alors dans le dénominateur de la dernière formule le carré de la masse de la particule chargée, qui est une quantité importante.

Dans le cas où l'onde incidente est plane, la densité de flux de l'énergie électromagnétique est numériquement égale au vecteur de Poynting (1852-1914)

$$S = \frac{c}{4\pi}EH = \frac{c}{4\pi}E_0^2\cos^2\omega t.$$

Sa moyenne temporelle est $\bar{S} = \frac{c}{8\pi}E_0^2$. En divisant l'énergie dissipée moyenne par $\bar{S}$, on trouve la *section de diffusion totale* par un électron libre

$$\sigma_T = \frac{8\pi}{3}\frac{e^4}{m_e^2c^4} = 0{,}6652448\,(33)\cdot 10^{-24}\ \text{cm}^2. \qquad (10.1)$$

Cette formule fut déjà établie par Thomson tout au début de l'élaboration de la théorie électronique. La grandeur σ_T est appelée *section de diffusion de Thomson des électrons*. On peut l'écrire sous la forme

$$\sigma_T = \frac{8}{3}\pi r_e^2, \qquad (10.2)$$

où r_e est le *rayon classique de l'électron* :

$$r_e = \frac{e^2}{m_ec^2} = 2{,}8179380\,(70)\cdot 10^{-13}\ \text{cm}. \qquad (10.3)$$

D'après (10.1), dans la gamme des rayons X, l'intensité de la radiation diffusée est indépendante de la fréquence de la radiation

incidente. Par contre, dans la gamme optique, l'intensité de la lumière diffusée par les atomes et les molécules, ainsi que par toutes les petites inhomogénéités du milieu, est proportionnelle à la quatrième puissance de la fréquence (cf. t. IV, § 98). Cette différence tient à ce que dans la gamme des rayons X, l'accélération de l'électron (dont dépend la diffusion) est déterminée par le champ électrique lui-même, tandis que dans la gamme optique, les dimensions des atomes et des molécules étant petites devant la longueur d'onde, l'intensité du champ électrique ne détermine que le moment dipolaire de la particule $p = \beta E$ (si on néglige les effets de résonance). Or la diffusion est proportionnelle au carré de sa dérivée seconde par rapport au temps, i.e. à la quatrième puissance de la fréquence ω. La même chose vaut pour les inhomogénéités du milieu, à condition que leurs dimensions linéaires soient petites devant λ.

La formule exacte de Klein-Nishina pour la section totale de la diffusion de Compton par un électron libre immobile est

$$\frac{\sigma}{\sigma_T}=\frac{3}{4}\left\{\frac{1+\gamma}{\gamma^3}\left[\frac{2\gamma(1+\gamma)}{1+2\gamma}-\ln(1+2\gamma)\right]+\frac{1}{2\gamma}\ln(1+2\gamma)-\frac{1+3\gamma}{(1+2\gamma)^2}\right\}, \tag{10.4}$$

où $\gamma = h\nu/m_ec^2$ est le rapport de l'énergie du quantum incident à l'énergie de l'électron au repos. Cette formule est en bon accord avec les résultats expérimentaux. A la différence de la formule de Thomson (10.3), la formule (10.4) prévoit une diminution monotone de σ à mesure qu'augmente l'énergie des quanta incidents. Cette décroissance de σ est illustrée par les données du tableau 1. La valeur maximale de σ, égale à σ_T, correspond à $\gamma \to 0$.

Tableau 1

Section droite de la diffusion de Compton en fonction de l'énergie des photons primaires

γ	0,05	0,1	0,2	0,33	0,5	1	2	3
σ/σ_T	0,913	0,84	0,737	0,637	0,563	0,431	0,314	0,254

γ	5	10	20	50	100	200	500	1000	
σ/σ_T	19,1	12,3	7,54	3,76	2,15	1,22	0,556	0,304	$\times 10^{-2}$

3. Considérons maintenant la diffusion d'un faisceau incident de rayons X par l'atome tout entier. Dans le cas d'un rayonnement X

dur, les électrons de l'atome diffuseront ce rayonnement *indépendamment* les uns des autres, donc de façon *incohérente*. Cette conclusion s'applique *a fortiori* aux électrons d'atomes différents. Considérons maintenant un faisceau parallèle de rayons X, ayant une section droite d'aire unité, se propageant suivant l'axe X. Entre deux sections droites du faisceau de coordonnées x et $x + dx$, on trouve $n\,dx$ atomes, n étant le nombre d'atomes contenus dans l'unité de volume. Un atome neutre contient Z électrons et ces électrons diffusent, dans l'unité de temps, une énergie égale à $Z\sigma I$; les électrons des $n\,dx$ atomes diffuseront donc une énergie égale à $Z\sigma I dx$, où I désigne l'intensité du faisceau incident. Par suite de la diffusion, l'intensité du faisceau diminue de $I\varkappa\,dx$, $\varkappa$ étant le *coefficient de diffusion* défini par la relation

$$\varkappa = \sigma n Z. \tag{10.5}$$

L'affaiblissement du faisceau résulte non seulement de la diffusion mais encore de l'*absorption* des rayons X. L'absorption s'accompagne d'un dégagement de chaleur à l'intérieur du corps et peut donc, ne serait-ce qu'en principe, être séparée de la diffusion. Le coefficient de diffusion $\varkappa$ est proportionnel à n, donc à la densité ρ du corps. Il s'ensuit qu'il convient de mesurer le rapport $\varkappa/\rho$. On sait que $\rho = nAm_{\text{H}}$, où m_{H} est la masse de l'atome d'hydrogène, A la masse atomique relative de la substance diffusante. En prenant pour σ la valeur de Thomson (10.1), on trouve

$$\frac{\varkappa}{\rho} = \frac{\sigma_{\text{T}}}{m_{\text{H}}}\frac{Z}{A} \approx 0{,}40\,\frac{Z}{A}\ \text{cm}^2\cdot\text{g}^{-1}. \tag{10.6}$$

On a constaté que pour les éléments légers les valeurs expérimentales de $\varkappa/\rho$ sont indépendantes de la longueur d'onde et voisines de 0,20 cm^2/g. Avec cette valeur on déduit de (10.6) que $Z/A \approx 1/2$, ce qui signifie que pour les éléments légers (l'hydrogène excepté), le nombre de charge Z (numériquement égal au numéro atomique de l'élément) est approximativement égal à la moitié du nombre de masse A. C'est ce que l'on observe pour les éléments situés au début de la classification périodique des éléments. La cause physique de cette régularité, ainsi que les écarts existants, seront expliqués lorsqu'on étudiera la structure du noyau atomique.

§ 11. Régularités spectrales

1. Les corps solides incandescents émettent des spectres de rayonnement *continus*. Dans le cas de corps gazeux, on observe, à côté d'une région à spectre continu, des *spectres de raies* et des *spectres de bandes*. Un spectre de raies est formé d'une suite de raies plus ou moins fines, réparties d'une certaine façon. Dans un spectre de bandes constitué par une série de cannelures, celles-ci présentent un aspect

continu lorsqu'on les observe à l'aide d'un spectroscope à faible pouvoir de résolution. Lorsqu'on utilise un instrument à haut pouvoir de résolution, on constate que ces bandes sont formées de nombreuses raies très rapprochées.

Au début du XXe siècle, on a établi que *les spectres de raies des gaz sont émis par des atomes ou des ions* et que *les spectres de bandes sont émis par des molécules.* C'est pour cela qu'on les appelle aussi *spectres atomiques* et *spectres moléculaires.* On n'arrive à observer le spectre atomique de l'hydrogène, dans une décharge électrique amorcée dans un tube à vide rempli d'hydrogène, que lorsque la majorité des molécules d'hydrogène auront été dissociées en atomes. Dans la vapeur d'iode, les bandes du spectre moléculaire disparaissent presque complètement lors de la dissociation des molécules I_2 en atomes I.

L'existence d'un grand nombre de raies spectrales reflète une structure interne complexe de l'atome étudié. Il n'est donc pas étonnant que l'énorme masse de données factuelles accumulées de façon empirique lors de l'étude des spectres constitua, au XXe siècle, la base sur laquelle fut fondée la théorie de la structure atomique.

On caractérise la position d'une raie spectrale dans le spectre par la *longueur d'onde* λ ou par la *fréquence* $\nu = c/\lambda$. L'utilisation de la fréquence est mieux adaptée pour mettre en évidence les régularités spectrales. Or, pour calculer les fréquences, il faut connaître la vitesse de la lumière c, dont la valeur était insuffisamment précise, tout au moins jusqu'à ces derniers temps. La longueur d'onde λ peut être mesurée avec précision, jusqu'à la septième décimale, si ce n'est au-delà, à l'aide d'appareils spectraux adéquats. C'est pour cette raison que les spectroscopistes utilisent à la place de la fréquence ν le *nombre d'onde spectroscopique* $\bar{\nu}$. C'est le nombre d'ondes contenues dans le vide dans un intervalle de 1 cm de longueur:

$$\bar{\nu} = \frac{1}{\lambda}\ *). \qquad (11.1)$$

En spectroscopie, on dénote la fréquence et le nombre d'onde spectroscopique par la même lettre ν. Pour éviter toute confusion, on dénotera la fréquence par ν et le nombre d'onde spectroscopique

*) Actuellement, tous ces développements ont perdu leur intérêt. En utilisant les procédés de l'optique non linéaire, on a réussi à mesurer la fréquence des vibrations lumineuses avec une précision meilleure que celle des mesures spectroscopiques des longueurs d'onde. Comme la vitesse de la lumière dans le vide est liée à la fréquence ν et à la longueur d'onde λ par la relation $c = \lambda\nu$, l'Assemblée Générale des poids et mesures adopta, en octobre 1983, une nouvelle définition du mètre. Selon cette définition, la vitesse de la lumière dans le vide est posée *exactement* égale à $c = 2{,}99792458 \cdot 10^8$ m/s. On définit le mètre par la distance que la lumière parcourt dans le vide en une seconde. Par conséquent, il est en principe indifférent d'utiliser en spectroscopie la quantité $1/\lambda$ ou la quantité c/λ.

par le symbole $\bar{\nu}$ (lettre ν surmontée d'un trait). D'ailleurs, il est préférable d'utiliser la notation $1/\lambda$ à la place de $\bar{\nu}$.

2. La loi fondamentale de la spectroscopie, qui fut établie empiriquement en 1908, est le *principe de combinaison de Ritz*. Selon ce principe, n'importe quelle raie spectrale d'un atome donné peut être déterminée en combinant deux par deux un nombre beaucoup plus petit (que le nombre de raies du spectre) de grandeurs appelées *termes spectraux* ou *termes* tout court. (Les termes sont désignés par T.) La fréquence (ou le nombre d'onde) de chaque raie spectrale s'exprime par la différence de deux termes:

$$\bar{\nu} = \frac{1}{\lambda} = T_{n_1} - T_{n_2}. \tag{11.2}$$

Par convention, on pose que les termes sont des grandeurs essentiellement positives et on les numérote de telle façon que l'accroissement du numéro du terme corresponde à une diminution de sa grandeur. Par exemple, dans la formule ci-dessus, on doit avoir $n_1 < n_2$ et $T_{n_1} > T_{n_2}$. Si on fixe la valeur de n_1 et si on attribue à n_2 des valeurs croissantes en partant de la valeur $n_2 = n_1 + 1$, on obtient un système de raies appelé *série spectrale*. L'ensemble des séries spectrales ainsi définies constitue le *spectre* (atomique) de l'élément considéré.

Soient deux raies spectrales d'une même série:

$$\bar{\nu}_{12} = T_{n_1} - T_{n_2}, \quad \bar{\nu}_{13} = T_{n_1} - T_{n_3}.$$

On admet que $\bar{\nu}_{12} > \bar{\nu}_{13}$ (donc $n_2 > n_3$). En soustrayant la deuxième égalité de la première, il vient

$$\bar{\nu}_{12} - \bar{\nu}_{13} = T_{n_3} - T_{n_2}.$$

Or c'est le nombre d'onde d'une autre raie spectrale du même élément chimique appartenant à une série commençant par le terme T_{n_3}. Ainsi, selon le principe de combinaison, *en calculant la différence des fréquences (ou des nombres d'onde) de deux raies spectrales appartenant à une seule et même série, on obtient la fréquence (le nombre d'onde) d'une raie spectrale appartenant à une autre série du même atome.* Il peut d'ailleurs arriver que la raie ainsi calculée n'existe pas dans le spectre de l'atome, car les combinaisons mutuelles des termes sont soumises à des limitations appelées *règles de sélection.*

3. On ne dispose pas d'expressions analytiques pour les termes de la grande majorité d'éléments. Dans les cas les plus favorables, on dispose de formules approchées empiriques ou semi-empiriques de certains termes. On indiquera au § 34 les formules des termes des atomes des métaux alcalins. La seule exception connue concerne l'atome le plus simple de tous, i.e. l'atome d'hydrogène composé d'un seul proton et d'un seul électron. Pour l'atome d'hydrogène, les

termes spectraux sont donnés avec une grande précision par la formule

$$T_n = \frac{R_H}{n^2} \quad (n = 1, 2, 3, \ldots), \tag{11.3}$$

où R_H est la *constante de Rydberg* (1854-1919) pour l'hydrogène, dont la valeur numérique est

$$R_H = 109678{,}76\ (1)\ \mathrm{cm}^{-1}. \tag{11.4}$$

Une expression semblable, où la constante de Rydberg possède une valeur différente de (11.4), s'applique à tous les isotopes d'hydrogène et à tous les ions monoélectroniques (cf. § 27).

Partant de l'expression (11.3), on établit par combinaisons les séries spectrales suivantes:

Série de Lyman:

$$\bar{\nu} = \frac{1}{\lambda} = R_H\left(1 - \frac{1}{n^2}\right) \quad (n = 2, 3, 4, \ldots). \tag{11.5}$$

Cette série fut découverte en 1916 par Lyman (1874-1954) dans la région ultraviolette du spectre.

Série de Balmer:

$$\bar{\nu} = \frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad (n = 3, 4, 5, \ldots). \tag{11.6}$$

Les quatre premières raies de cette série se trouvent dans la région visible du spectre et sont notées H_α, H_β, H_γ, H_δ ; toutes les autres raies se situent dans l'ultraviolet. Balmer (1825-1898) réussit à formuler la règle (11.6) en se fondant sur les positions relatives des quatre premières raies, ce qui marqua le commencement d'une systéma-

Tableau 2

Longueurs d'onde de la série de Balmer de l'hydrogène

n	Notation	λ, nm	
		calculée	mesurée
3	H_α	656,279	656,285
4	H_β	486,133	486,132
5	H_γ	434,047	434,046
6	H_δ	410,174	410,173
7	H_ε	397,008	397,007
8	H_ζ	388,905	388,906
9	H_η	383,539	383,540
10	H_ϑ	379,790	379,791
11	H_ι	377,067	377,065

tisation des séries spectrales. Une représentation schématique de la série de Balmer est donnée sur la figure 20. D'autre part, le tableau 2, où sont rassemblées les valeurs calculées et les valeurs expérimentales des longueurs d'onde des 9 premières raies de la série de Balmer, montre l'excellent accord entre les valeurs calculées à l'aide de la

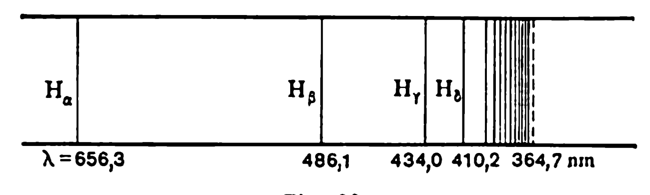


Fig. 20

formule (11.6) et les résultats expérimentaux. Comme les longueurs d'onde ont été mesurées dans l'air, les longueurs d'onde calculées à l'aide de la formule (11.6) ont été corrigées de façon convenable.

Série de Paschen (1865-1947):

$$\bar{\nu} = \frac{1}{\lambda} = R_H \left(\frac{1}{3^2} - \frac{1}{n^2}\right) \quad (n = 4, 5, 6). \tag{11.7}$$

Cette série fut prévue en 1908 par Ritz par application du principe de combinaison. Toutes les raies de cette série se déduisent en combinant les raies de la série de Balmer. La même année. Paschen l'observa dans la région infrarouge du spectre.

Série de Brackett:

$$\bar{\nu} = \frac{1}{\lambda} = R_H \left(\frac{1}{4^2} - \frac{1}{n^2}\right) \quad (n = 5, 6, 7, \ldots). \tag{11.8}$$

Série de Pfund (1879-1949):

$$\bar{\nu} = \frac{1}{\lambda} = R_H \left(\frac{1}{5^2} - \frac{1}{n^2}\right) \quad (n = 6, 7, 8, \ldots). \tag{11.9}$$

Ces deux dernières séries, qui se situent dans l'infrarouge lointain, ont été découvertes, la première en 1922, la deuxième en 1924. La série de Brackett se déduit par combinaisons des raies de la série de Paschen, et la série de Pfund se déduit de la série de Brackett.

Dans la série de Lyman, la plus grande longueur d'onde correspond à $n = 2$ et a pour valeur $\lambda = \frac{4}{3R} = 121,56713$ nm. La raie correspondante est appelée *raie de résonance de l'hydrogène*. Pour trouver la fréquence (ou le nombre d'onde) maximale, on applique les formules (11.5) à (11.9) en posant $n = \infty$. Cette fréquence est appelée *fréquence limite de la série*. Par exemple, pour la série de Balmer, la limite est

$$\bar{\nu}_\infty = R_H/4 = 27419,69 \text{ cm}^{-1} \quad \text{ou} \quad \lambda_\infty = \frac{4}{R_H} = 364,70142 \text{ nm}.$$

A mesure qu'on se rapproche de la limite d'une série, les raies spectrales se resserrent, i.e. la différence des longueurs d'onde des raies adjacentes tend asymptotiquement vers zéro; les intensités des raies tendent également vers zéro. Au-delà de la limite d'une série, on trouve un spectre continu. Cette règle s'applique aux séries spectrales de l'hydrogène et d'autres éléments chimiques. Il y existe les limites des séries au-delà desquelles apparaît un spectre continu.

§ 12. Les postulats de Bohr

1. Les lois de la physique classique sont, par essence, adaptées à la description des *processus continus*. Or la finesse des raies spectrales émises par les atomes des éléments chimiques témoigne de ce que les processus évoluant à l'intérieur des atomes présentent un caractère *discontinu* ou *discret*. A côté du caractère continu, le caractère discret doit être reflété par les lois fondamentales de la physique. Bohr (1885-1962) fut le premier à formuler, en 1913, deux postulats qui répondaient à ce nouveau besoin.

Postulat n° 1. *Un atome (ou un système atomique) ne peut pas se trouver dans tous les états prévus par la mécanique classique; seuls certains états quantiques, caractérisés par des valeurs discrètes bien déterminées de l'énergie* $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$, ..., *peuvent être occupés par les systèmes atomiques. En contradiction avec les prévisions de l'électrodynamique classique, un atome occupant ces états, dits stationnaires, n'émet aucun rayonnement.*

La mécanique quantique, qui a supplanté la théorie de Bohr, conduit tout naturellement au concept d'états stationnaires caractérisés par des niveaux d'énergie discrets; il est vrai qu'elle admet aussi l'existence d'états stationnaires présentant des niveaux d'énergie continus, mais dans ce cas les électrons et les noyaux ne forment plus un système lié, et certaines de ces particules exécutent des mouvements infinis. Dans les atomes et les molécules, les particules qui les constituent sont liées et exécutent donc des mouvements finis. Dans ce dernier cas, les niveaux d'énergie sont discrets, conformément au premier postulat de Bohr (cf. § 13,3 ainsi que § 22). Nous avons déjà mentionné plus haut qu'en mécanique quantique le postulat de Bohr n'est pas une proposition indépendante et n'est qu'un corollaire des principes fondamentaux de la mécanique quantique.

Bohr admit initialement que le mouvement des électrons dans les états stationnaires d'un atome pouvait être décrit par les mêmes notions et les mêmes grandeurs que celles utilisées en mécanique classique pour caractériser le mouvement des particules. Autrement dit, il admettait que les électrons se mouvaient sur des trajectoires bien définies et qu'à tout moment leur mouvement pouvait être caractérisé par des valeurs déterminées des coordonnées et des vitesses. Or

ce point de vue est incompatible avec les principes fondamentaux de la mécanique quantique, qui rejette le concept de trajectoires sur lesquelles se meuvent les électrons et les autres microparticules (cf. § 20). C'est pour cela que tout en admettant le premier postulat de Bohr, nous caractériserons les états stationnaires seulement par les valeurs de l'énergie sans essayer de visualiser d'aucune façon le mouvement des électrons dans les atomes.

Postulat n° 2. *Lors de la transition d'un état stationnaire de plus grande énergie $\mathcal{E}_{n_2}$ à un état stationnaire d'énergie plus petite $\mathcal{E}_{n_1}$, l'énergie de l'atome varie de $\mathcal{E}_{n_2} - \mathcal{E}_{n_1}$. Si cette variation d'énergie est due à l'émission d'un rayonnement, ce processus s'accompagne de l'émission d'un photon d'énergie*

$$h\nu \equiv \hbar\omega = \mathcal{E}_{n_2} - \mathcal{E}_{n_1}. \tag{12.1}$$

La même relation vaut pour l'absorption lorsqu'un photon incident fait passer un atome sur un niveau de plus grande énergie et disparaît lui-même. La relation (12.1) exprime la *règle des fréquences de Bohr.*

2. Ainsi, un système atomique passe d'un état stationnaire à un autre par *sauts.* Ces sauts sont dits *quantiques.* La théorie de Bohr ne disait pas ce que devenait le système pendant un saut, i.e. entre deux états stationnaires. Certains physiciens éminents estimaient que cette question était, par principe, dénuée de sens, n'étant pas susceptible de vérification expérimentale. Ces savants s'appuyaient sur le principe épistémologique, selon lequel la science ne doit pas chercher à décrire la nature à l'aide de considérations ne pouvant être vérifiées par l'expérience. La science ne doit avoir affaire qu'à des grandeurs observables ne serait-ce qu'en principe. Formulée de cette façon, cette exigence est dénuée de sens et ne fournit aucun élément nouveau. Il en serait tout autrement si cette proposition était complétée par un critère de ce qui est observable en principe et de ce qui ne l'est pas. Or on ne peut établir un tel critère *a priori. Seule la théorie* pourrait apporter une réponse à cette question, la justesse de la théorie devant être vérifiée par l'expérience. Les théories existantes se distinguent, entre autres, par la réponse qu'elles donnent à cette même question. La théorie de Bohr, se refusant à expliquer ce qui se passe pendant un saut quantique, témoigne de ses insuffisances. Cependant, si une théorie se refusait à s'occuper de phénomènes concrets, tant qu'elle n'aurait pas répondu à la question considérée, elle serait assurément inutile. Dans le cas qui nous occupe, la situation est exactement la même que dans le cas de la théorie de la gravitation universelle. Newton arriva à ses fins parce qu'il s'appliqua à tirer des conclusions concrètes de la loi qu'il découvrit, laissant ouvert le problème de la nature physique de la gravitation.

3. Les processus faisant intervenir plusieurs photons sont également possibles, cela lorsque la transition d'un niveau à un autre s'accompagne de l'émission ou de l'absorption de plusieurs photons si-

multanément. Mais ce sont des processus *non linéaires* qui ne peuvent se produire qu'en présence de champs forts, comme ceux qui apparaissent dans les impulsions géantes de rayonnement laser. Nous ne les considérerons pas ici.

Il existe enfin des *transitions non radiatives* où l'énergie est libérée ou absorbée non pas sous forme de rayonnement mais sous forme de chaleur. Un atome peut transiter d'un niveau à un autre non pas à la suite de l'émission ou de l'absorption d'un photon, mais, par exemple, à la suite d'un choc avec une autre particule.

4. La règle des fréquences de Bohr (12.1) permet d'expliquer le principe de combinaison de Ritz. L'application de la règle de Bohr au calcul du nombre d'onde spectroscopique $\bar{\nu} = \nu/c$ conduit au résultat suivant:

$$\bar{\nu} = 1/\lambda = \frac{\mathcal{E}_{n_2}}{ch} - \frac{\mathcal{E}_{n_1}}{ch}. \qquad (12.2)$$

En comparant ce résultat à la relation (11.2), on obtient

$$T_n = -\frac{\mathcal{E}_n}{ch}. \qquad (12.3)$$

Ainsi se trouve révélée la signification physique des termes spectraux: ils sont déterminés par les niveaux énergétiques de l'atome.

Un autre résultat expérimental se laisse expliquer par la règle de Bohr. Si on observe dans le spectre d'émission de l'hydrogène atomique la série le Lyman, on observe nécessairement toutes les autres séries spectrales — celles de Balmer, de Paschen, etc. Par contre, on n'observe dans le spectre d'absorption de l'hydrogène atomique non lumineux que la série de Lyman, toutes les autres séries étant absentes. La raison en est que pour faire apparaître dans le spectre d'émission la série de Lyman, il faut fournir une énergie suffisante pour arracher un électron occupant le niveau d'énergie le plus bas. Puisqu'il faut une énergie moindre pour exciter les niveaux qui se trouvent plus haut dans l'échelle des énergies, les électrons qui les occupent seront libérés. Ainsi des transitions entre une multitude de niveaux deviennent possibles et on verra apparaître dans le spectre d'émission différentes séries spectrales. La situation est différente dans le cas des spectres d'absorption. Si l'hydrogène ne rayonne pas, c'est que ses atomes se trouvent dans l'état normal, i.e. occupent le niveau le plus bas, et il n'y a aucun autre niveau occupé par des atomes. Lorsque de la lumière est absorbée, seules des transitions entre le niveau le plus bas et les niveaux supérieurs sont possibles. Ces transitions ne font apparaître dans le spectre d'absorption que la série de Lyman. Si on veut faire apparaître d'autres séries de raies dans le spectre d'absorption, il faut qu'avant éclairement d'autres

niveaux énergétiques soient déjà excités. Mais dans ces conditions, apparaîtra le spectre d'émission du gaz considéré. Les raies d'absorption de ces séries ne peuvent donc être observées que superposées au spectre d'émission du gaz. Il faut que ces dernières soient moins lumineuses que les raies correspondantes du rayonnement incident. Les spectres atomiques des autres éléments se comportent de façon analogue.

5. L'ensemble des valeurs de l'énergie des états stationnaires de l'atome $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$, ... constituent le *spectre énergétique de l'atome*. Selon la théorie classique, les fréquences émises par un atome (dans le cas simple du mouvement unidimensionnel) forment également une suite ν_1, ν_2, ν_3, ..., à condition de les numéroter à l'aide d'un seul indice. Certains physiciens estiment qu'il y a là une certaine contradiction entre la théorie classique et les résultats expérimentaux. L'expérience conduit effectivement au principe de combinaison de Ritz. Selon ce principe et la théorie de Bohr, chaque fréquence résulte de la combinaison de deux termes et doit donc être numérotée à l'aide de *deux indices*, par exemple ν_{12}, ν_{13}, ... ; le premier indice désigne le numéro du terme initial, le second désignant le numéro du terme final. Par conséquent, les fréquences doivent former non pas une suite unidimensionnelle, mais un tableau rectangulaire ou une *matrice*. Cette objection est le résultat d'un malentendu et nous ne la signalons que parce qu'on la trouve dans certains ouvrages, et un lecteur non averti pourrait la prendre au sérieux. Or les mathématiques enseignent que les éléments d'une matrice rectangulaire forment un ensemble fini ou dénombrable et on peut donc les numéroter à l'aide d'*un seul indice*, même si le nombre d'éléments contenus dans chaque ligne et dans chaque colonne de la matrice est infiniment grand.

§ 13. Spectre de l'hydrogène

1. On appelle *quantification*, ou plus exactement *quantification de l'énergie* d'un atome, la détermination des valeurs $\mathcal{E}_1$, $\mathcal{E}_2$, $\mathcal{E}_3$, ... de l'énergie des états stationnaires d'un atome. Bohr proposa une règle de quantification pour l'atome d'hydrogène, qui fournissait des résultats corrects. Sous une forme générale, le problème de la quantification a été formulé par la mécanique quantique non seulement pour l'atome d'hydrogène mais aussi pour *tout* système atomique (cf. § 22). Il est assez compliqué. La règle de quantification proposée par Bohr ne présente qu'un intérêt historique. Néanmoins il nous paraît utile de donner ici une solution simple du problème de la quantification de l'atome d'hydrogène ou de l'atome hydrogénoïde, proche des idées de Bohr. Cette solution se fonde sur une analogie avec la mécanique classique et sur une expression empirique établie pour les termes spectraux de l'atome d'hydrogène.

Posons que les termes spectraux et les niveaux d'énergie correspondants de l'atome d'hydrogène sont de la *forme de Balmer*

$$T_n = \frac{Z^2R}{n^2}, \tag{13.1}$$

$$\mathscr{E}_n = -chT_n = -\frac{chZ^2R}{n^2}, \tag{13.2}$$

où R est la *constante de Rydberg*, le nombre de charge Z du noyau a été introduit pour raison de commodité. Le nombre entier n est appelé *nombre quantique principal*. A mesure que n augmente, les niveaux énergétiques adjacents de l'atome se rapprochent, et pour $n \to \infty$ la distance entre les niveaux tend vers zéro. Ainsi le caractère discret du spectre énergétique s'estompe de plus en plus à mesure que n croît. On peut donc s'attendre à ce qu'à la limite le système quantique se comportera comme un système classique. Cette proposition est due à Bohr, qui la dénomma *principe de correspondance*.

L'application du principe de correspondance permet d'exprimer la constante de Rydberg R à l'aide des constantes fondamentales caractérisant l'atome. Pour plus de généralité, on considérera un *atome hydrogénoïde* ; on désigne par ce terme un ion dont le noyau porte une charge égale à $+Ze$, autour duquel gravite un seul électron. Pour $Z = 1$, on obtient un atome d'hydrogène neutre H, pour $Z = 2$, on obtient un atome d'hélium une fois ionisé He^+, pour $Z = 3$, on obtient un atome de lithium deux fois ionisé Li^{++}, etc. Il va de soi que les développements qui suivent présupposent la validité de l'expression (13.2) pour tout atome hydrogénoïde. La constante de Rydberg est-elle la même pour différents atomes hydrogénoïdes ? C'est ce que nous verrons plus tard.

2. Pour simplifier, Bohr supposa que l'électron tournait autour du noyau sur une orbite circulaire. Plus tard, Sommerfeld (1868-1951) généralisa les résultats de Bohr au cas d'orbites elliptiques. Cette généralisation est devenue inutile après l'avènement de la mécanique quantique et nous ne l'exposerons pas. On peut donc se limiter au cas simple d'orbites circulaires. Selon les conceptions classiques, la fréquence de la lumière émise est égale à la fréquence de rotation de l'électron sur une orbite circulaire. Aux basses fréquences, ce résultat est indubitable, comme le démontre la comparaison entre les résultats de la théorie classique et les données expérimentales dans la gamme des ondes radioélectriques. Les calculs qui suivent concernent ces fréquences-là. On doit obtenir une concordance des résultats que fournissent la théorie quantique et la théorie classique pour que le principe de correspondance soit vérifié. On admettra que le noyau est infiniment lourd, ce qui revient à poser qu'il est immobile. Lorsqu'un électron tourne sur une orbite circu-

laire de rayon r avec une pulsation ω

$$m\omega^2 r = \frac{Ze^2}{r^2}, \tag{13.3}$$

d'où $\omega = Ze^2/(Lr)$, où $L = mr^2\omega$ est le moment cinétique de l'électron. L'énergie totale de l'électron se compose de ses énergies potentielle et cinétique:

$$\mathcal{E} = \frac{1}{2} mr^2\omega^2 - \frac{Ze^2}{r} = -\frac{Ze^2}{2r}.$$

Selon la théorie classique, on doit donc avoir

$$\omega = -\frac{2\mathcal{E}}{L}. \tag{13.4}$$

D'autre part, les niveaux d'énergie d'un atome hydrogénoïde doivent être de la forme de Balmer (13.2). Il s'ensuit que lors des transitions d'un atome entre les niveaux, la quantité $\mathcal{E}_n n^2$ doit se conserver: $\mathcal{E}_n n^2 =$ const. Pour les grands nombres quantiques n et pour les petites variations Δn de ces nombres, on doit avoir

$$\frac{\Delta\mathcal{E}}{\mathcal{E}_n} + 2\frac{\Delta n}{n} = 0.$$

Compte tenu de la règle des fréquences de Bohr $\Delta\mathcal{E} = \hbar\omega$, on obtient

$$\omega = -\frac{2\mathcal{E}}{\hbar n}\Delta n. \tag{13.5}$$

On a supprimé l'indice n auprès de $\mathcal{E}$ et on pose $\Delta n > 0$ pour ne pas avoir affaire à des fréquences négatives. La fréquence la plus basse correspond à la transition $\Delta n = 1$; c'est la *fréquence fondamentale*. Aux valeurs $\Delta n = 2, 3, \ldots$ correspondent différents *harmoniques* de la fréquence fondamentale. D'après le principe de correspondance, la fréquence fondamentale figurant dans (13.5) doit coïncider avec la fréquence classique (13.4). Pour cela il faut que

$$L = n\hbar. \tag{13.6}$$

Il s'ensuit que d'après la théorie de Bohr, le moment cinétique (pour n grand) est quantifié, i.e. ne peut prendre que des valeurs discrètes définies par (13.6).

De la formule (13.3) on peut tirer maintenant

$$(mr^2\omega)^2 = Ze^2rm = (n\hbar)^2,$$

d'où

$$r = \frac{\hbar^2}{Ze^2m}n^2 \tag{13.7}$$

et par suite

$$\mathscr{E}_n = -\frac{Ze^2}{2r} = -\frac{(Ze^2)^2 m}{2\hbar^2 n^2}. \qquad (13.8)$$

En identifiant cette dernière formule à (13.2), on obtient

$$R_\infty = \frac{me^4}{4\pi c\hbar^3} = \frac{2\pi^2 me^4}{ch^3} = 109737{,}309 \pm 0{,}12 \text{ cm}. \qquad (13.9)$$

On a affecté R de l'indice ∞ afin de spécifier que la formule (13.9) a été établie en posant que la masse M du noyau est infinie et le noyau immobile. Dans cette approximation, la constante de Rydberg a la même valeur pour tous les atomes hydrogénoïdes.

Quoique la valeur théorique (13.9) de la constante de Rydberg soit très proche de la valeur expérimentale trouvée pour l'atome d'hydrogène $R_H = 109677{,}576$ cm^{-1}, la différence entre ces deux valeurs est très grande, compte tenu de la très grande précision des mesures spectroscopiques. Cette différence tient à ce qu'on n'a pas tenu compte de la nature finie de la masse M du noyau dans les développements conduisant à la formule (13.9). Pour en tenir compte, il faut remplacer la masse m de l'électron par la masse réduite $Mm/(M + m)$. On obtient alors

$$R = \frac{R_\infty}{1 + m/M}. \qquad (13.10)$$

Dans cette approximation, la constante de Rydberg dépend de la masse du noyau et par conséquent ses valeurs, pour les différents atomes hydrogénoïdes, doivent différer les unes des autres, quoique cette différence soit très petite. Appliquée à l'atome de l'hydrogène, la formule (13.10) fournit la valeur $R_H = 109677{,}6$ cm^{-1}, qui est en bon accord avec la valeur expérimentale.

La formule (13.10) peut servir au calcul de la constante de Rydberg R_∞ pour le cas d'un noyau infiniment lourd. Pour ce faire, il suffit de disposer de la valeur spectroscopique de R, pour l'hydrogène par exemple, et de la valeur du rapport m/M donnée par les mesures spectroscopiques de masse.

3. On a établi la formule (13.8) pour les grandes valeurs du nombre quantique n; mais cette formule reste valable *pour n'importe quelles valeurs de n*, puisqu'on a postulé que les termes vérifiaient la formule de Balmer (13.1), qui n'impose aucune limitation à la valeur de n. Nous avions pour objectif d'établir, sur la base du principe de correspondance, les formules théoriques (13.9) et (13.10) pour les constantes de Rydberg R_∞ et R. Comme ces constantes ne sauraient dépendre de n, on a pu déterminer leurs valeurs à l'aide de calculs où n était supposé grand.

En spectroscopie, il est d'usage de représenter les termes spectraux et les niveaux d'énergie par des lignes horizontales, et les transitions entre ces derniers par des flèches. Les flèches orientées

d'un niveau supérieur vers un niveau inférieur correspondent à des raies d'émission et celles qui pointent en sens inverse correspondent à des raies d'absorption. A titre d'exemple, sur la figure 21 on a représenté de cette façon le spectre de l'hydrogène. Les niveaux d'énergie y sont numérotés par le nombre quantique n, le zéro de

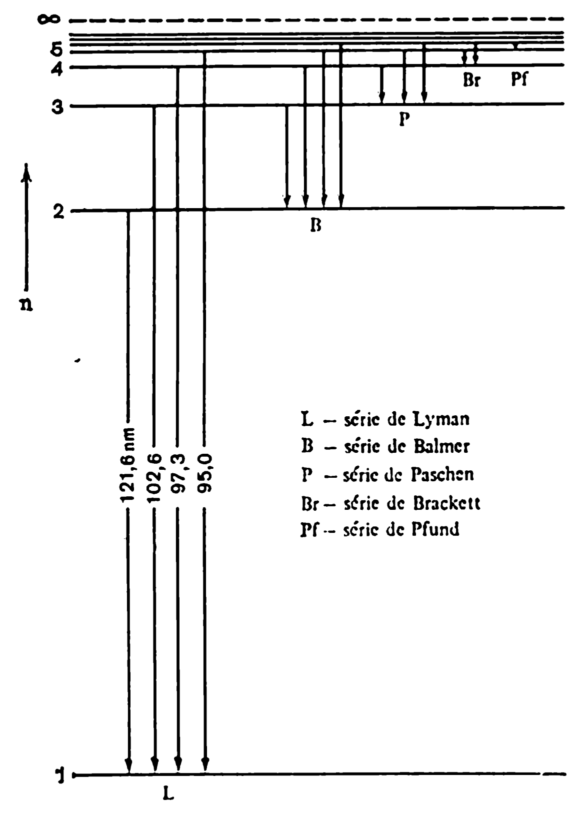


Fig. 21

l'échelle correspond à l'énergie du niveau correspondant à $n = \infty$; ce niveau est resprésenté par la droite horizontale en pointillé située tout en haut. Tous les niveaux d'énergie situés en dessous sont discrets; à ces niveaux correspondent des valeurs négatives de l'énergie totale de l'atome. Au-dessus de la ligne en pointillé, l'énergie n'est plus quantifiée, ce qui signifie que le spectre énergétique y est continu. Lorsque $\mathscr{E} < 0$, le mouvement de l'électron est fini et lorsque $\mathscr{E} > 0$, il est infini; cela résulte directement d'un théorème correspondant de la mécanique classique (cf. t. I, §§ 25, 57) que

l'on peut appliquer ici pour de grands n. Ce théorème peut être démontré en mécanique quantique conséquente (§ 22), i.e. d'une façon parfaitement rigoureuse.

Nous voyons ainsi que *le noyau et l'électron ne forment un système lié, donc un atome, que si le spectre énergétique est discret.* Dans le cas d'un spectre énergétique continu, l'électron peut s'éloigner du noyau aussi loin que possible; le couple noyau-électron ne peut alors être appelé atome que tout conventionnellement. Si on n'entend par atome que les états liés, on peut dire que les niveaux d'énergie de l'atome sont nécessairement discrets, comme le postula Bohr.

L'existence d'électrons non liés rend cependant possibles les transitions quantiques entre les états du spectre énergétique continu, ainsi qu'entre ces états et les états du spectre d'énergie discret. Ces différentes transitions correspondent à l'apparition d'un spectre d'émission ou d'absorption continu qui est superposé au spectre de raies de l'atome. C'est ce qui explique que le spectre atomique ne s'arrête pas à la limite d'une série mais s'étend vers les ondes plus courtes où il devient continu. Les transitions quantiques depuis des états du spectre énergétique continu (i. e. des états où l'atome est ionisé) dans des états du spectre discret s'accompagnent de la recombinaison des électrons avec les ions positifs correspondants. Le rayonnement résultant de ces transitions est appelé *rayonnement de recombinaison.*

Le passage d'un atome de son état normal à un niveau énergétique plus élevé du spectre discret correspond à une *excitation de l'atome.* Le passage d'un atome d'un des niveaux du spectre discret dans la région du spectre continu transforme l'atome en un système non lié. C'est le processus d'*ionisation de l'atome.* Si l'atome se trouvait initialement dans son état normal, l'énergie d'ionisation minimale serait: $\mathcal{E}_{\text{ion}} = \mathcal{E}_\infty - \mathcal{E}_1 = -\mathcal{E}_1$, ce qui donne pour un atome hydrogénoïde

$$\mathcal{E}_{\text{ion}} = \frac{(Ze^2)^2 m}{2\hbar^2} = 2\pi c\hbar Z^2 R_\infty = chZ^2 R_\infty \qquad (13.11)$$

ou

$$\mathcal{E}_{\text{ion}} = 21{,}8\cdot 10^{-12} Z^2 \text{ ergs} = 13{,}6\ Z^2 \text{ eV}. \qquad (13.12)$$

4. Jusqu'en 1913, i.e. avant que fut connue la théorie de Bohr, on attribuait à l'hydrogène encore deux séries:

$$\frac{1}{\lambda} = R_{\text{H}}\left(\frac{1}{1{,}5^2} - \frac{1}{\bar{n}^2}\right), \quad \bar{n} = 2, 3, 4, \ldots, \qquad (13.13)$$

$$\frac{1}{\lambda} = R_{\text{H}}\left(\frac{1}{2^2} - \frac{1}{(\bar{n}+1/2)^2}\right), \quad \bar{n} = 2, 3, 4, \ldots \qquad (13.14)$$

La série (13.13) fut découverte par Fowler (1889-1944) dans un mélange (H + He), tandis que la série (13.14) fut observée par Picke-

ring dans le spectre de la nébuleuse planétaire (ζ-PUPPS). Or, selon la théorie de Bohr, les raies de ces séries correspondent non pas à l'hydrogène H, mais à l'hélium ionisé He^+. Ces raies font partie des séries du spectre de He^+:

$$\frac{1}{\lambda}=4R_{He}\left(\frac{1}{3^2}-\frac{1}{n^2}\right),\quad n=4,\ 5,\ 6,\ \ldots, \tag{13.13a}$$

$$\frac{1}{\lambda}=4R_{He}\left(\frac{1}{4^2}-\frac{1}{n^2}\right),\quad n=5,\ 6,\ 7\ \ldots \tag{13.14a}$$

On a négligé ici la différence entre les valeurs des constantes de Rydberg pour l'hydrogène et l'hélium. En fait ces constantes diffèrent légèrement l'une de l'autre, comme le montre la formule

$$\frac{R_{He}}{R_H}=\frac{1+m_e/M_H}{1+m_e/M_{He}}=1{,}00041. \tag{13.15}$$

En négligeant cette petite différence entre R_{He} et R_H, les raies de la série (13.13) se superposent dans cette approximation aux raies de la série (13.13a) pour n pairs, tandis que pour les valeurs impaires de n ce sont les raies des séries (13.14) et (13.14a) qui se superposent. Effectivement, les séries (13.13a) et (13.14a) ont été décelées dans le spectre de l'hélium pur. Assurément, il n'y a aucune raison de diviser les séries entières (13.13a) et (13.14a) en moitiés pour obtenir les séries de Fowler (13.13) et de Pickering (13.14). Comme R_{He} est un peu plus grand que R_H, les raies spectrales de l'hélium ionisé He^+ sont légèrement déplacées du côté des ondes courtes par rapport aux raies de la série de Balmer pour l'hydrogène. C'est l'effet dit de *déplacement isotopique des raies spectrales.*

Bien entendu, cet effet existe aussi pour les autres éléments chimiques et leurs ions, quoique pour les atomes à plusieurs électrons on ne puisse l'interpréter aussi facilement que pour les atomes hydrogénoïdes. En effet, les différents isotopes d'un même élément chimique ou d'un même ion ne se distinguent que par les masses de leurs noyaux; ils portent les mêmes charges et possèdent donc les mêmes enveloppes électroniques. Or les processus d'émission de lumière se produisent dans les *couches électroniques.* L'expression « déplacement isotopique » des raies de l'hélium par rapport aux raies de l'hydrogène semble être illogique, car l'hélium n'est pas l'isotope d'hydrogène. Cependant, dans l'effet de déplacement des raies spectrales il s'agit non pas d'atomes d'hélium mais de ses *atomes ionisés une fois.* Quant aux ces derniers, ils se comportent dans ce cas en tant qu'isotopes d'hydrogène. Pour les autres éléments la situation est la même.

On notera cependant que le déplacement isotopique des raies spectrales résulte non seulement de la différence des masses des isotopes, mais aussi des dimensions différentes de leurs noyaux atomiques. Anticipant sur ce qui suit, on dira que cette conclusion s'im-

pose aussitôt si on fonde la quantification sur l'équation de Schrödinger (cf. § 22). En effet, les champs de force coulombiens et nucléaires régnant dans les noyaux doivent être quelque peu différents pour les isotopes. Comme les régions de l'espace occupées par ces champs sont également différentes, les fonctions d'onde et les valeurs propres correspondantes de l'énergie doivent différer également suivant les isotopes. L'influence des dimensions des noyaux sur le déplacement isotopique des raies spectrales est particulièrement notable dans le cas de noyaux lourds. Dans ce cas, le déplacement isotopique dû aux dimensions des noyaux est du même ordre de grandeur que le déplacement isotopique dû à la seule différence des masses.

5. On sait qu'il existe outre l'hydrogène ordinaire deux isotopes d'hydrogène : le *deutérium* et le *tritium*. Le noyau de l'atome d'hydrogène ordinaire se compose d'un seul proton. L'hydrogène ordinaire est appelé *protium*. Le noyau de l'atome de deutérium est appelé *deuton* et celui de l'atome de tritium, *triton*. Un deuton est formé par un proton et un neutron, et un triton comporte un proton et deux neutrons. Ces noyaux sont respectivement 2 et 3 fois plus lourds que le noyau du protium. L'atome de deutérium est dénoté par D ou ^2H, et l'atome de tritium par T ou ^3H. Le deutérium est un élément stable, tandis que le tritium est un élément radioactif. Le tritium est produit artificiellement en bombardant une cible en lithium ou en béryllium avec des deutons rapides. Il se forme aussi en quantité infime dans l'atmosphère, sous l'action des rayons cosmiques. Dans les eaux naturelles, pour 6800 atomes de protium on trouve en moyenne 1 atome de deutérium. Par suite d'une différence relativement grande des masses de D et de H, la différence de leurs propriétés physiques est plus marquée que chez les autres isotopes. Ainsi la densité de l'eau lourde D_2O est près de 10 % plus grande que celle de l'eau ordinaire H_2O ; la température de fusion de D_2O est égale à 3,8 °C (sous pression atmosphérique normale). Le déplacement isotopique des raies spectrales pour D et H n'est guère important. Ainsi, par exemple,

$$\frac{R_D}{R_H}=\frac{1+m_e/M_H}{1+m_e/M_D}=1{,}000272.$$

Il s'ensuit que le déplacement isotopique des fréquences ou des longueurs d'onde des raies du deutérium par rapport aux raies de l'hydrogène n'est que de

$$\frac{\Delta\nu}{\nu}=-\frac{\Delta\lambda}{\lambda}=\frac{R_D-R_H}{R_H}=m_e\left(\frac{1}{M_H}-\frac{1}{M_D}\right)=0{,}000272.$$

La raie D_α de la série de Balmer du deutérium est décalée par rapport à la raie H_α de l'hydrogène du côté des ondes courtes de $|\Delta\lambda| =$ $= 0{,}179$ nm.

6. Les calculs donnés ci-dessus ne sont justifiés que pour les grands nombres quantiques n. Or Bohr admit initialement que dans les états stationnaires les électrons se mouvaient sur des orbites bien déterminées, comme le font les particules en mécanique classique. Dans ce cas, même pour les petites valeurs de n, doivent être valables aussi bien la formule (13.8) pour l'énergie de l'état stationnaire que la formule (13.7) donnant le rayon de l'orbite de l'électron. La mécanique quantique réfuta les orbites, mais la formule (13.7) y conserve sa signification puisqu'elle donne l'ordre de grandeur des dimensions de l'atome dans les états stationnaires correspondants. La signification exacte de cette expression sera précisée au § 27 et ici on notera que le rayon de l'orbite de l'électron dans l'atome d'hydrogène ordinaire, pris dans son état normal (i.e. pour $n = 1$), est appelé *rayon de Bohr* que l'on note r_B. Selon (13.7), il est égal à

$$r_B = \frac{\hbar^2}{me^2} = 0,52917 \cdot 10^{-8} \text{ cm}. \tag{13.16}$$

C'est l'ordre de grandeur des dimensions atomiques que donnait jadis la théorie cinétique des substances. L'intensité du champ électrique créé par le noyau, régnant sur la première orbite de Bohr de l'atome d'hydrogène, est

$$E = \frac{e}{r_B^2} = 17,2 \cdot 10^6 \text{ CGSE} = 51,5 \cdot 10^8 \text{ V/cm}.$$

D'une façon générale, la valeur de $E \sim 10^8$ V/cm est caractéristique de l'échelle des intensités des champs électriques intra-atomiques. Dans un champ électrique extérieur ayant une intensité de cet ordre, les atomes sont rapidement ionisés.

On établit aisément que la vitesse avec laquelle un électron se meut sur une orbite circulaire stationnaire est donnée par l'expression

$$v = \frac{Z}{n} \alpha c, \tag{13.17}$$

où α est une constante sans dimension:

$$\alpha = \frac{e^2}{\hbar c} = 1/137,03604(11) \approx 1/137, \tag{13.18}$$

que l'on appelle *constante de la structure fine*. Pour le mouvement sur la première orbite de Bohr de l'atome d'hydrogène,

$$v = \alpha c = \frac{c}{137} = 0,00730c. \tag{13.19}$$

En introduisant la constante α dans la formule (13.8), on obtient

$$\mathscr{E}_n = -\frac{Z^2}{2n^2} \alpha^2 mc^2, \tag{13.20}$$

i.e. $\mathscr{E}_n$ est exprimé en fonction de l'énergie propre mc^2 de l'électron.

PROBLÈMES

1. Déterminer les frontières de la région spectrale où sont disposées les raies de la série de Balmer de l'hydrogène atomique.

Réponse. $\lambda_{B\infty} < \lambda < \lambda_{BC}$, où $\lambda_{B\infty} = 4/R_H = 364{,}705$ nm,

$$\lambda_{BC} = 36/(5R_H) = 656{,}468 \text{ nm (raie rouge)}.$$

2. En utilisant les résultats du problème précédent, calculer les longueurs d'onde des frontières côté ondes courtes des séries de Lyman et de Paschen de l'hydrogène.

Réponse. Pour la série de Lyman, $\lambda_{L\infty} = \lambda_{B\infty}/4 = 91{,}1762$ nm. Pour la série de Paschen, $\lambda_{P\infty} = {}^{9}/_{4}\lambda_{B\infty} = 820{,}586$ nm.

3. Calculer l'énergie qu'il faut communiquer à un atome d'hydrogène pour que sa série de Balmer ne contienne qu'une seule raie spectrale.

Réponse. $\mathcal{E}_1 < \mathcal{E} < \mathcal{E}_2$, où $\mathcal{E}_1 = {}^{8}/_{9}hcR_H = 12{,}09$ eV,

$$\mathcal{E}_2 = {}^{15}/_{16}hcR_H = 12{,}75 \text{ eV}.$$

4. Quelles raies spectrales verra-t-on apparaître dans le spectre de l'hydrogène atomique lorsqu'il sera éclairé par des radiations ultraviolettes de 100 nm de longueur d'onde?

Réponse. $\lambda_1 = {}^{4}/_{3}R_H = 121{,}5682$ nm,

$$\lambda_2 = {}^{9}/_{8}R_H = 102{,}57317 \text{ nm}, \quad \lambda_3 = {}^{36}/_{5}R_H = 656{,}46828 \text{ nm}.$$

5. Un atome d'hydrogène peut-il absorber des radiations de nombre d'onde $1/\lambda = 3R$?

Réponse. L'absorption se produit et donne lieu à l'ionisation de l'atome d'hydrogène.

6. Un atome d'hydrogène initialement immobile émet un photon dont la fréquence correspond à celle de la première raie de la série de Lyman. Calculer la variation relative de la fréquence $\Delta\nu/\nu_0$ du photon résultant du recul de l'atome. Quelle sera la vitesse v de l'atome après émission du photon?

Réponse. $\dfrac{\nu_0 - \nu}{\nu_0} = \dfrac{h\nu_0}{2m_{at}c^2} = \dfrac{3}{16}\dfrac{m_e}{m_{at}}\alpha^2 = 5{,}44 \cdot 10^{-9}$;

$$v = \frac{3}{8}\frac{m_e}{m_{at}}\alpha^2 c = 326 \text{ cm/s},$$

où $\alpha = e^2/\hbar c = 1/137$ est la constante de la structure fine.

7. Calculer les potentiels d'ionisation des ions He^+ et Li^{++}.

Réponse. $V_{He} = 4V_H = 54$ V; $V_{Li} = 9V_H = 122$ V.

8. Trouver la longueur d'onde de la raie de résonance de He^+.

Réponse. $\lambda = 30{,}375$ nm.

9. Calculer l'énergie minimale qu'il faut communiquer à un atome de béryllium trois fois ionisé, pris dans son état fondamental, pour exciter son spectre complet.

Réponse. $\mathcal{E}_{Be} = hcZ^2R_H = 217{,}5747$ eV.

10. Un photon de la tête de la série de Lyman de l'ion hélium He^+ est absorbé par un atome d'hydrogène pris à l'état fondamental et provoque son ionisation. Calculer l'énergie cinétique $\mathcal{E}_{cin}$ qu'acquerra l'électron lors de l'ionisation.

Réponse. $\mathcal{E}_{cin} = 2\,\mathcal{E}_{H,ion} = 27{,}2$ eV, où $\mathcal{E}_{H,ion}$ est l'énergie d'ionisation de l'atome d'hydrogène.

11. Evaluer à l'aide de la théorie de Bohr le nombre de raies spectrales que l'on peut faire apparaître dans différentes séries du spectre d'émission de l'hydrogène atomique, contenu dans un tube à décharge sous une pression $P = 5$ mm Hg, à la température $T = 300$ K.

Solution. Les séries de Lyman, Balmer, Paschen, etc., de l'hydrogène n'apparaissent que si les atomes émetteurs sont isolés, c'est-à-dire si le diamètre

de l'orbite de Bohr $2r = 2n^2r_B$ ne dépasse pas la distance moyenne entre les atomes $l = (kT/P)^{1/3}$. Partant de là, on peut évaluer le numéro maximal admissible n de l'orbite par la formule $n \approx \sqrt{l/(2r_B)}$. Le plus grand nombre de raies que l'on pourra observer dans le tube à la pression et à la température indiquées sera égal à

dans la série de Lyman: $n - 1 \approx 12$,
dans la série de Balmer: $n - 2 \approx 11$,
dans la série de Paschen: $n - 3 \approx 10$.

12. Dans les spectres d'émission de certaines étoiles, on dénombre $m \sim 40$ raies de la série de Balmer de l'hydrogène. Quel doit être le nombre minimal N des traits d'un réseau de diffraction pour assurer la résolution de ces raies dans un spectre de premier ordre?

Réponse. $N \approx m^3/8 = 8000$ traits.

13. A proximité de la raie spectrale de l'hydrogène $\lambda_1 = 486,1320$ nm, Urey (1893-1981) décela, en 1932, une nouvelle raie $\lambda_2 = 485,9975$ nm. En supposant que l'apparition de cette dernière raie résulte de la présence dans l'hydrogène ordinaire d'une petite quantité de l'un de ses isotopes, déterminer la masse atomique relative m_x/m_H de cet isotope.

Solution. A l'aide de la formule (13.10), on obtient

$$\frac{\lambda_1}{\lambda_2} \approx \frac{1-m_e/m_H}{1-m_e/m_x},$$

où m_x est la masse atomique de l'isotope inconnu. En négligeant la quantité m_e/m_x devant l'unité, on en tire

$$\frac{\lambda_1-\lambda_2}{\lambda} \approx \frac{m_e}{m_x}-\frac{m_e}{m_H}.$$

Par conséquent

$$\frac{m_e}{m_x}=\frac{m_e}{m_H}-\frac{\lambda_1-\lambda_2}{\lambda}.$$

Comme $m_e/m_H = 1/1835$, on en déduit $m_e/m_x = 1/3727$ et par suite $m_x/m_H \approx 2$. La raie λ_2 est due au deutérium.

14. Calculer la différence entre les longueurs d'onde des séries de Balmer pour l'hydrogène et le deutérium dans la partie visible du spectre.

Réponse.

$$\lambda_{H_\alpha}-\lambda_{D_\alpha}=0,179 \text{ nm},$$
$$\lambda_{H_\beta}-\lambda_{D_\beta}=0,132 \text{ nm},$$
$$\lambda_{H_\gamma}-\lambda_{D_\gamma}=0,118 \text{ nm},$$
$$\lambda_{H_\delta}-\lambda_{D_\delta}=0,116 \text{ nm}.$$

15. Quel doit être le pouvoir de résolution R (ne pas confondre avec la constante de Rydberg) d'un appareil spectral pour pouvoir observer le déplacement isotopique des raies spectrales du deutérium par rapport aux raies de l'hydrogène. Quelle doit être la largeur b de la base d'un prisme en flint lourd ayant une dispersion $dn/d\lambda = 1000$ cm^{-1} (dans la gamme des radiations rouges) utilisé dans un spectrographe à prisme, afin de pouvoir déceler le déplacement isotopique de la raie de tête (de grande longueur d'onde) de la série de Balmer?

Réponse.

$$R > 2\frac{m_H}{m_e} \approx 3700; \quad b > \frac{R}{dn/d\lambda} \approx 3,7 \text{ cm}.$$

16. On veut observer la série de Lyman d'un mélange d'hydrogène et de tritium atomiques. Quel doit être le pouvoir de résolution R du spectroscope pour séparer la structure isotopique des raies spectrales de cette série? Quel devrait être le pouvoir de résolution pour séparer la structure isotopique des raies des autres séries (de Balmer, Paschen, etc.) de ce même mélange? Peut-on séparer la structure isotopique des raies spectrales de ce mélange dans la région visible du spectre en utilisant un prisme de verre dont l'épaisseur à la base est $b = 1$ cm et la dispersion de l'indice de réfraction $dn/d\lambda = 1000$ cm^{-1}? Quels doivent être le nombre effectif N_{eff} de réflexions et l'ordre m du spectre observé pour séparer et étudier cette même structure isotopique à l'aide de l'interféromètre de Fabry-Pérot?

R é p o n s e. Le pouvoir de résolution doit être supérieur à $^3/_2 m_p/m_e \approx$ ≈ 2800. Il est le même pour toutes les raies des séries spectrales du mélange. Le pouvoir de résolution du prisme $b\, dn/d\lambda = 1000$, donc insuffisant pour séparer les raies. $m < {}^3/_2 m_p/m_e < N_{eff} m$.

17. Connaissant les valeurs de la constante de Rydberg pour l'hydrogène et l'hélium ($R_{He} = 109722{,}267$ cm^{-1}) ainsi que R_∞, calculer la valeur du rapport m_{He}/m_H de la masse de l'atome d'hélium à la masse de l'atome d'hydrogène.

R é p o n s e.

$$m_{He}/m_H \approx \frac{R_\infty - R_H}{R_\infty - R_{He}} = 3{,}971.$$

18. La charge massique e/m de l'électron peut être calculée par un procédé spectroscopique. Effectuer ce calcul en utilisant les données spectroscopiques

$$R_H = 109677{,}576 \text{ cm}^{-1}, \quad R_D = 109707{,}419 \text{ cm}^{-1},$$

les masses relatives des moles d'hydrogène et de deutérium atomiques

$$H = 1{,}008142, \quad D = 2{,}014755$$

(la masse de l'atome est égale à la somme des masses du noyau et de l'électron) et le nombre de Faraday

$$F = N_{Av} e = 2{,}89366 \cdot 10^{14} \text{ CGSE/mole} = 9632{,}19 \text{ CGSM/mole}.$$

S o l u t i o n. Les constantes de Rydberg pour l'hydrogène et le deutérium sont données par les expressions

$$R_H = \frac{R_\infty}{1 + m_e/M_H}, \quad R_D = \frac{R_\infty}{1 + m_e/M_D},$$

M_H et M_D étant les masses des noyaux d'hydrogène et de deutérium. On en tire

$$\frac{1}{m_e} = \frac{1}{R_D - R_H}\left(\frac{R_H}{M_H} - \frac{R_D}{M_D}\right).$$

En multipliant les deux membres de cette relation par e et en remarquant que $(M_H + m_e) N_{Av} = H$, $(M_D + m_e) N_{Av} = D$, $N_{Av} e = F$, il vient

$$\frac{e}{m_e} = \frac{F}{R_D - R_H}\left(\frac{R_H}{H - N_{Av} m_e} - \frac{R_D}{D - N_{Av} m_e}\right).$$

Comme la masse de l'électron est petite devant la masse de l'atome, il n'est pas nécessaire de calculer $N_{Av} m_e$ de façon précise. Il suffit de prendre $N_{Av} m_e =$ $= 6{,}022 \cdot 10^{23} \cdot 9{,}11 \cdot 10^{-28} = 5{,}49 \cdot 10^{-4}$; $H - N_{Av} m_e = 1{,}007593$; $D - N_{Av} m_e =$ $= 2{,}014186$. En substituant ces valeurs dans la formule précédente, on trouve

$$e/m_e = 5{,}2732 \cdot 10^{17} \text{ CGSE/g} = 1{,}7590 \cdot 10^7 \text{ CGSM/g}.$$

La précision de ce calcul se perd lorsqu'on détermine la différence $R_D - R_H$. Il faut donc connaître les valeurs de R_H et R_D avec une très grande précision. A noter que la constante de Rydberg R_∞ ne figure pas dans le calcul.

19. Le positonium (symbole chimique ps) est le système lié formé d'un électron et d'un positon, tournant autour du centre de masse de ce système. Le positonium se forme lors du choc de positons lents avec des atomes; pendant le choc le positon capte un électron atomique. C'est un système instable, car l'électron et le positon s'annihilent mutuellement en donnant naissance à deux (en $\sim 10^{-10}$ s) ou à trois (en $\sim 10^{-7}$ s) quanta gamma. Trouver pour le positonium les niveaux d'énergie, l'énergie d'ionisation et la longueur d'onde de la raie de résonance.

Réponse.

$$\mathscr{E}_{ps} = -\frac{hc}{2n^2} R_\infty; \qquad \mathscr{E}_{ion} = \frac{hc}{2} R_\infty = 6{,}80 \text{ eV},$$

$$\lambda_{res} = 243 \text{ nm}.$$

20. Lorsque des particules élémentaires chargées négativement (muons μ^-, mésons π^-, mésons K^-, etc.) sont freinées dans une substance, elles peuvent être captées par le champ coulombien du noyau et se substituer aux électrons de l'enveloppe électronique des atomes. Le cas le plus intéressant est celui de la capture des muons négatifs. En pratique, on n'arrive à remplacer qu'un seul électron, et les systèmes substitués qui en résultent sont appelés *mésatomes*. La masse du muon est $m_\mu = 207 m_e$. Calculer par la théorie de Bohr le rayon de la première orbite circulaire (orbite K) du muon dans le mésatome. Calculer ses niveaux énergétiques. Quel rayonnement est émis lorsqu'un muon passe d'une orbite supérieure sur l'orbite K ? Pourquoi l'étude de ce rayonnement est utilisée pour identifier la structure superficielle des noyaux atomiques lourds? On négligera la masse du muon devant celle du noyau.

Solution. L'enveloppe électronique de l'atome n'affecte pratiquement pas le mouvement du muon dans le mésatome pour la raison que le muon se trouve très près du noyau ou dans le noyau lui-même. Par raison de la symétrie sphérique du nuage électronique, on peut poser que le champ électrique créé par ce dernier est nul à l'emplacement du muon. Si l'orbite du muon se trouve hors du noyau, le rayon de l'orbite K du muon est

$$r_1 = \frac{h^2}{4\pi^2 e^2 m_\mu Z} = \frac{0{,}0026}{Z} \cdot 10^{-8} \text{ cm},$$

cette valeur est près de 200 fois plus petite que la valeur de r_1 pour un atome hydrogénoïde de même nombre de charge Z. Pour ce même cas, les niveaux énergétiques sont définis par l'expression

$$\mathscr{E}_n = -\frac{hcR_\infty Z^2 m_\mu}{m_e} = -2820 Z^2 \text{ eV}.$$

Cette expression montre que le rayonnement, émis lorsqu'un muon transite sur l'orbite K venant d'orbites d'énergie plus grande, se situe dans la région X du spectre et, pour les grandes valeurs de Z, dans la région des rayons gamma. Lorsque Z est grand, l'orbite K du muon se trouve à l'intérieur du noyau de l'atome et les formules ci-dessus cessent d'être valables. Les résultats dépendent fortement de la distribution de la charge électrique dans le noyau et c'est ce qui permet d'utiliser les mésatomes pour l'étude de la structure des noyaux atomiques.

§ 14. Preuves expérimentales des postulats de Bohr

1. Les résultats présentés dans le paragraphe précédent constituent une preuve expérimentale solide des postulats de Bohr. Ces postulats sont à la base des conceptions physiques qui ont été mises en

œuvre lors de la création et des perfectionnements des générateurs quantiques de la gamme optique appelés lasers (cf. t. IV, ch. XI). Les lasers ont ouvert des voies parfaitement nouvelles pour la physique, les techniques et les autres domaines de la science. Le fait que les lasers fonctionnent conformément aux prévisions de la théorie justifie la théorie et, partant, les postulats de Bohr.

Dans ce paragraphe, on présente d'autres résultats expérimentaux qui confirment eux aussi les postulats de Bohr. Examinons en premier lieu les expériences de James Franck (1882-1964) et de Gustav Hertz (1887-1975) qui ont été entreprises peu avant la parution de

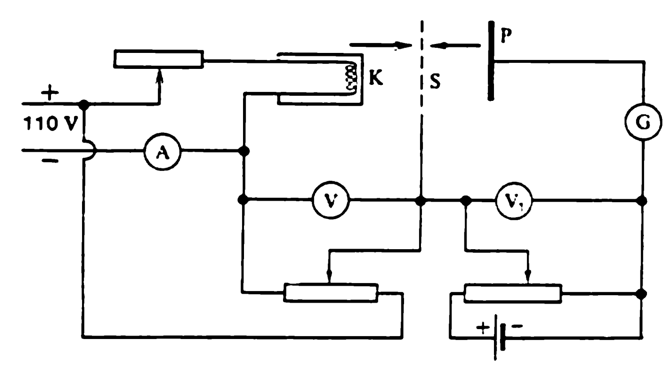


Fig. 22

la théorie de Bohr. Les expériences avaient pour objectif initial la mesure des *potentiels d'ionisation des atomes*. Mais ces expériences ont apporté une preuve de la justesse des postulats de Bohr. Dans ces expériences, Franck et Hertz faisaient passer des électrons convenablement accélérés à travers le gaz étudié. Lorsque les atomes étaient heurtés par les électrons, ils passaient dans des *états excités*, chacun de ces états étant caractérisé par une énergie déterminée. Si les niveaux énergétiques de l'atome sont discrets, comme l'affirme le premier postulat de Bohr, l'énergie cinétique des électrons doit être égale ou supérieure à une certaine valeur minimale pour qu'ils commencent à exciter les atomes du gaz. Ce procédé d'excitation des atomes convient particulièrement bien aux gaz inertes et aux vapeurs métalliques dont les atomes sont dénués d'affinité électronique, i.e. ne manifestent aucune aptitude à la capture des électrons et à la formation d'ions négatifs. Dans leurs premières expériences, Franck et Hertz travaillaient avec de la vapeur de mercure. Bien entendu, la vapeur de mercure doit être relativement dense pour que les chocs entre les électrons et les atomes soient assez fréquents. Dans les expériences, on faisait varier la densité de la vapeur de mercure en disposant le tube contenant quelques gouttes de mercure dans un four dont on ré-

glait la température qu'on maintenait constante pendant la durée de l'expérience.

2. Le schéma du dispositif expérimental de Franck et Hertz est représenté sur la figure 22. La différence de potentiel V, servant à accélérer les électrons, est appliquée entre la cathode chaude K et la grille S. Les électrons émis par K traversent la vapeur de mercure et parviennent à la grille S. S'ils possèdent encore une énergie cinétique suffisante, ils passeront au travers de la grille et après avoir surmonté un petit potentiel de freinage V_1 (de l'ordre de 0,5 V) ils tomberont sur le collecteur P. Si l'énergie cinétique des électrons est insuffisante, ils ne pourront pas surmonter le poteniel de freinage

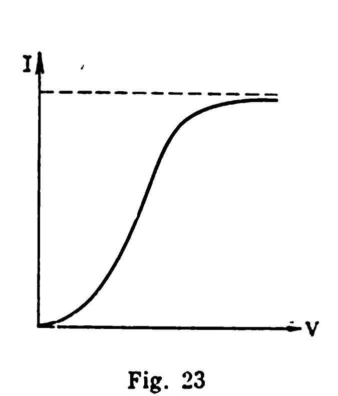


Fig. 23

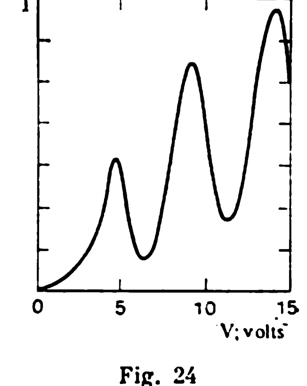


Fig. 24

V_1 et parvenir au collecteur. Le courant électronique I passant par le collecteur était mesuré par le galvanomètre G. L'ampèremètre A servait à contrôler la constance du courant de chauffage de la cathode. Le ballon de verre où se trouvent la cathode, la grille et le collecteur n'est pas représenté sur la figure 22.

Si on avait fait le vide dans le ballon, la dépendance du courant thermo-ionique I avec la tension d'accélération V serait représentée par la courbe de la figure 23. Par suite de l'existence d'une charge d'espace, la partie initiale de cette courbe correspondrait à la loi des 3/2. Pour les grandes valeurs de la tension appliquée, on arriverait à la saturation où I ne dépend plus de V (cf. t. III, § 101). Les expériences ont montré qu'en présence de la vapeur de mercure, la courbe $I = I(V)$ a une allure qualitativement différente, puisqu'on y voit apparaître plusieurs maximums et minimums de courant. Pour le mercure, l'intervalle entre deux minimums ou deux maximums voisins a été trouvé égal à 4,9 eV (fig. 24). L'existence de ces maximums et minimums est liée au caractère discret des niveaux énergétiques de l'atome; pour l'atome de mercure le niveau le plus proche du niveau

normal (c'est le niveau le plus bas) se trouve à une distance de 4,9 eV environ.

3. Les chocs entre les électrons et les atomes de mercure peuvent être *élastiques* ou *inélastiques*. Dans le premier cas, l'état interne de l'atome de mercure ne change pas et, par conséquent, l'énergie cinétique de l'électron n'est pas utilisée pour exciter l'atome. Cette énergie ne peut être que transférée à l'atome de mercure dont l'énergie cinétique augmente. Mais il n'est pas nécessaire de tenir compte de ce processus puisque la masse de l'atome de mercure est près de 400 mille fois plus grande que la masse de l'électron. Lorsqu'un électron heurte un atome de mercure, il en est réfléchi comme s'il heurtait un mur et la direction de son mouvement est seule à changer. Il est inutile de tenir compte du mouvement thermique de l'atome puisque l'énergie moyenne de ce mouvement, égale à $^3/_2\ kT$, est très petite par rapport à l'énergie que possède un électron accéléré par un champ électrique. Tant que l'énergie cinétique de l'électron est inférieure à $\mathscr{E}_1 \approx 4{,}9$ eV, il ne peut pas exciter l'atome de mercure et tous les chocs entre ces particules seront élastiques. On conçoit que parmi les très nombreux chocs entre les électrons et les atomes de mercure, il en est qui déterminent un changement de la direction initiale des électrons, ainsi que d'autres qui rétablissent presque exactement la direction du mouvement initial. Si l'électron est en mesure de surmonter le potentiel de freinage avant choc, il pourra le faire aussi après plusieurs chocs élastiques, à condition que la direction de son mouvement n'aura pas changé. Par conséquent, tant que les chocs sont élastiques, le courant électronique I croît avec la tension appliquée V.

La situation devient autre lorsque l'énergie des électrons devient égale ou supérieure à $\mathscr{E}_1 = 4{,}9$ eV. L'électron peut alors dépenser la totalité ou une partie de son énergie cinétique pour *exciter l'atome de mercure*. On dit que les chocs sont alors *inélastiques*. Si l'énergie cinétique restante est trop petite pour que les électrons surmontent le potentiel de freinage V_1, ils ne pourront pas arriver jusqu'au collecteur P et, en conséquence, le courant électronique I diminue à mesure que la tension appliquée V augmente. Lorsque l'énergie des électrons devient égale à $\mathscr{E}_1$, le courant électronique I ne tombe pas brusquement à zéro parce que les électrons possèdent des vitesses différentes, ne serait-ce que du fait de la dispersion des vitesses d'éjection de la cathode. D'autre part, pour qu'un électron puisse surmonter le champ retardateur, il faut faire intervenir non pas sa vitesse totale mais la *composante longitudinale* de la vitesse, qui est parallèle au champ; la composante transversale de la vitesse ne joue aucun rôle. La composante longitudinale change à chaque choc élastique, car seule la vitesse totale se conserve. Ces différents facteurs lissent la courbe $I = I\ (V)$ et expliquent pourquoi les maximums de cette courbe se situent non pas là où l'énergie est égale à 4,9 eV, mais pour des énergies plus petites.

Si on fait croître encore la tension d'accélération, tout électron qui, après avoir subi un choc élastique, reste dans le champ accélérateur peut être accéléré à nouveau et acquérir une énergie suffisante pour exciter le premier niveau énergétique $\mathscr{E}_1$ d'un atome de mercure. Si l'électron subit un nouveau choc inélastique et perd une partie de son énergie, il peut être accéléré une nouvelle fois (si le potentiel V est grand) et acquérir l'énergie nécessaire pour exciter un atome, etc. Ces multiples excitations du niveau $\mathscr{E}_1$ déterminent l'apparition, sur la courbe $I = I(V)$, des maximums successifs pour des valeurs de l'énergie égales à $\mathscr{E}_1$, $2\mathscr{E}_1$, $3\mathscr{E}_1$, etc. L'existence d'une *différence de potentiel de contact* entre la cathode et la grille fausse les indications du voltmètre V et déplace la courbe $I = I(V)$ toute entière à gauche ou à droite. Mais on peut éliminer la différence de potentiel de contact en déterminant la valeur de $\mathscr{E}_1$ par mesure de l'intervalle séparant deux maximums successifs du courant $I = I(V)$, car cet intervalle en est indépendant.

4. Un choc inélastique avec excitation du premier niveau excitable $\mathscr{E}_1$ ne se produit pas nécessairement. L'électron peut continuer à accumuler de l'énergie dans le champ électrique jusqu'à ce qu'elle devienne suffisante pour exciter le deuxième $\mathscr{E}_2$, le troisième $\mathscr{E}_3$, etc., niveau énergétique. Les tensions d'accélération correspondant à ces niveaux d'énergie sont appelées *potentiels d'excitation* ou *potentiels critiques*. Ainsi, pour des valeurs proches de ces potentiels, on peut voir apparaître de nouveaux maximums sur la courbe $I = I(V)$. Dans les gaz d'assez grande densité, cela ne se réalise pratiquement pas et c'est ce qui se manifesta dans les premières expériences de Franck et Hertz (par exemple, à la température de 210 °C, la pression de la vapeur de mercure est égale à 24 mm Hg). La raison en est que pour acquérir l'énergie nécessaire à l'excitation, l'électron doit pouvoir parcourir dans le champ électrique accélérateur, sans subir de chocs inélastiques, une distance grande devant la longueur de libre parcours entre deux chocs inélastiques successifs. Dans les gaz denses, cette condition n'est pas réalisable. L'excitation des niveaux énergétiques successifs $\mathscr{E}_2$, $\mathscr{E}_3$, . . . est possible à des pressions plus basses et avec un appareillage plus perfectionné (voir ci-dessous sous-paragraphe 6).

5. En modifiant le schéma représenté sur la figure 22, on peut l'adapter à la mesure des *potentiels d'ionisation des atomes. Le potentiel d'ionisation est la tension d'accélération minimale nécessaire pour initier l'ionisation des atomes et la formation d'ions positifs.* On conçoit que pour rendre l'ionisation possible, les électrons doivent acquérir par accélération une énergie au moins égale à l'*énergie d'ionisation de l'atome.* Dans le cas du mercure, il faut pour cela que la température du four soit assez basse, afin que la pression de la vapeur de mercure soit égale à 0,5 mm Hg environ. La seule modification à apporter au schéma de montage (fig. 22) consiste à réunir

le pôle positif du voltmètre V_1 au pôle négatif (au lieu du pôle positif) du voltmètre V. Dans ce cas, le potentiel du collecteur sera de 0,5 V environ plus petit que le potentiel de la cathode, quelle que soit la valeur de la tension d'accélération de grille. Les électrons émis par la cathode ne parviendront plus à atteindre le collecteur. Aucun courant ne traversera le galvanomètre G tant qu'il n'y aura pas d'ions positifs, i.e. tant que le potentiel d'accélération de grille sera inférieur au potentiel d'ionisation. Dès que la tension d'accélération de grille deviendra égale au potentiel d'ionisation, les ions positifs apparaîtront et le galvanomètre sera traversé alors par un courant. La tension que le voltmètre V indiquera à ce moment représentera le potentiel d'ionisation du mercure. Il va de soi qu'il faudra corriger la lecture du voltmètre de la valeur de la différence de potentiel de contact entre cathode et grille, qui aura été déterminée lors des mesures précédentes du premier potentiel d'excitation du mercure.

6. Pour améliorer le pouvoir résolvant du dispositif et observer les maximums correspondant aux différents niveaux énergétiques du mercure, on peut, tout en conservant le schéma de montage de la figure 22, utiliser deux grilles au lieu d'une et leur appliquer une tension d'accélération de l'ordre de 0,1 V seulement. La tension d'accélération principale est appliquée entre la cathode et la première grille et la tension de freinage est appliquée entre la deuxième grille et le collecteur. Cette disposition est très avantageuse, car on ménage ainsi entre les grilles un espace important où *le champ électrique est presque inexistant*. C'est dans cet espace que les électrons, accélérés jusqu'au potentiel d'excitation ou au-delà, subissent la grande majorité des chocs inélastiques. La distance entre la cathode et la grille d'accélération doit être petite dans ce cas, afin que le nombre de chocs y soit relativement petit. Ce sont ces particularités qui rehaussent le pouvoir résolvant de l'appareil. Bien entendu, la pression du gaz contenu dans le ballon doit être petite (pour la vapeur de mercure, 5 à 6 mm Hg). En procédant de la façon indiquée, on a mis en évidence l'existence, chez l'atome de mercure, d'un deuxième potentiel d'excitation $V = 6,7$ V. Le troisième potentiel d'excitation $V = 10,4$ V est déjà un potentiel d'ionisation.

Hertz a mis au point une méthode de mesure des potentiels critiques plus sensible. Les électrons sortant de la cathode K et accélérés par le potentiel V pénètrent dans un espace sans champ, qui est limité par une grille cylindrique représentée en pointillé (en coupe) sur la figure 25. C'est dans cet espace que se produisent les chocs entre les électrons et les atomes du gaz. Si les chocs sont élastiques, les électrons sont déviés sans perte d'énergie et s'échappent à travers la surface latérale de la grille dans un faible champ retardateur créé par le potentiel $V_1 \approx 0,1$ V appliqué entre la grille et le collecteur P (concentrique à la grille). Les électrons possèdent une énergie suffisante pour surmonter V_1 et parvenir jusqu'au collecteur P. Lorsqu'à la suite d'une augmentation de la tension d'accélération V des chocs inélastiques commencent à se produire, l'énergie que posséderont les électrons après choc peut être insuffisante pour surmonter le

champ retardateur et les électrons ne pourront plus atteindre le collecteur P. Le courant passant dans le circuit du collecteur P est mesuré par le galvanomètre G. Ces mesures se font pour deux valeurs proches du potentiel de freinage, par exemple $V_1 = 0$ et $V_1 = 0{,}1$ V.

Tant que les chocs sont élastiques, les indications du galvanomètre seront pratiquement les mêmes pour ces deux valeurs du potentiel V_1. Mais lorsque la tension d'accélération V devient égale à l'un des potentiels critiques, des chocs inélastiques apparaissent, et pour $V_1 = 0{,}1$ V le courant passant par le galvanomètre G diminue brusquement. La différence des indications du galvanomètre G correspondant à $V_1 = 0$ et $V_1 = 0{,}1$ V peut servir de mesure du nombre d'électrons ayant une énergie inférieure à 0,1 eV, qui s'échappent de la grille; c'est

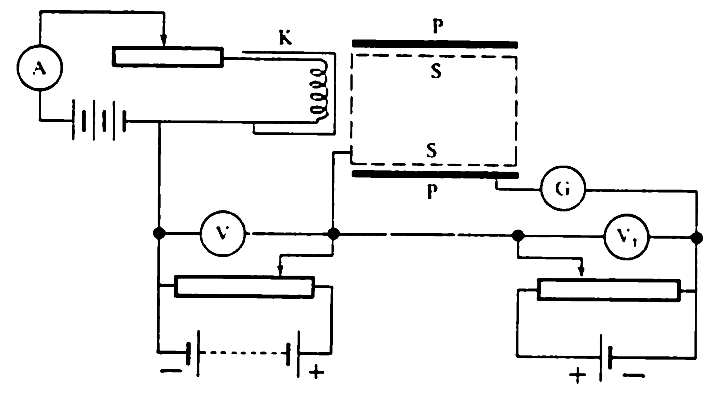


Fig. 25

le nombre d'électrons qui subissent des chocs inélastiques à proximité du potentiel critique considéré. Comme les chocs entre électrons et atomes se produisent dans une région où il n'y a pas de champ électrique, après les chocs les électrons ne peuvent plus acquérir l'énergie requise pour exciter les atomes. C'est pour cette raison qu'on ne voit plus apparaître sur la courbe $I = I(V)$ des maximums équidistants correspondant à un potentiel critique donné. Les maximums multiples qu'on pouvait observer avant n'existent plus et on ne peut exclure directement la différence de potentiel de contact. On peut cependant déterminer cette différence de potentiel de contact pour en tenir ultérieurement compte, si on connaît la valeur d'un potentiel critique au moins. On peut aussi utiliser pour cela un mélange de deux gaz, si on connaît les potentiels critiques de l'un d'eux.

7. Les expériences de Franck et Hertz ont apporté une confirmation de la justesse du second postulat de Bohr, à savoir de la règle des fréquences. Il s'avéra que la vapeur de mercure n'est pas lumineuse tant que la tension d'accélération est inférieure au premier potentiel critique $\approx 4{,}9$ V. Dans ce cas, aucun atome de mercure n'est excité et il ne peut donc y avoir de transitions entre les états excités et les états non excités. Lorsque la tension d'accélération atteint 4,9 V environ, les premiers atomes excités apparaissent. Les transitions de ces atomes dans des états non excités doivent provoquer une émission de lumière. Il se produit en effet l'émission d'une raie de résonance ultraviolette $\lambda = 253{,}7$ nm. Le ballon de verre où est

contenue la vapeur de mercure, ainsi que les autres pièces utilisées, ne permet pas d'observer cette radiation, le verre étant opaque aux rayons ultraviolets. Les pièces d'optique en verre ne conviennent pas non plus à l'étude de la composition spectrale du rayonnement émis. On devrait utiliser pour cela des matériaux comme le quartz ou la fluorite qui sont transparents à l'ultraviolet jusqu'aux longueurs d'onde de 180 et 120 nm respectivement.

Connaissant la longueur d'onde $\lambda = 253{,}7$ nm, on peut calculer le premier potentiel critique du mercure avec une précision bien supérieure à celle que permet l'écartement des maximums de la courbe $I = I(V)$ relevée dans les expériences de Franck et Hertz. On utilise alors la formule

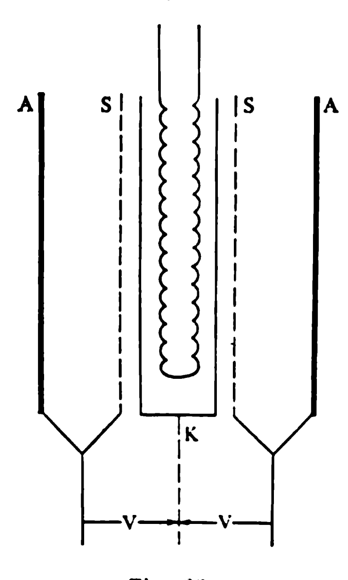


Fig. 26

$$\mathscr{E}_2 - \mathscr{E}_1 = eV = h\nu = \frac{hc}{\lambda}.$$

En y substituant les valeurs numériques, on trouve

$$\lambda\,(\text{nm}) = \frac{1240}{V\,(\text{V})}. \qquad (14.1)$$

En y posant $\lambda = 253{,}7$ nm, on trouve $V = 4{,}887$ V, résultat en bon accord avec la valeur de 4,9 V qui fut déterminée dans les expériences de Franck et Hertz.

La mise en œuvre d'une méthode expérimentale plus précise a permis de déterminer la valeur du deuxième potentiel critique $V = 6{,}7$ V. Les transitions de ce niveau dans l'état fondamental s'accompagnent de l'émission de radiations de longueur d'onde $\lambda = 184{,}9$ nm. En appliquant la formule (14.1), on voit qu'à cette raie spectrale correspond un potentiel $V = 6{,}705$ V, ce qui est en bon accord avec les résultats expérimentaux. Le troisième potentiel critique du mercure égal à 10,4 V correspond déjà à l'ionisation. Il s'ensuit que pour $V > 10{,}4$ V on doit observer des transitions de tous les états non liés sur tous les niveaux énergétiques situés plus bas dans l'échelle de l'énergie. L'expérience montre en effet que dans ces conditions on arrive à exciter tout le spectre atomique du mercure.

8. Pour exciter les atomes par des chocs électroniques et observer ensuite la luminosité du gaz, il est tout indiqué de modifier le montage expérimental, puisque le champ accélérateur devient alors inutile. Un schéma de principe convenant à l'observation du rayonnement émis par le gaz est représenté sur la figure 26. La tension d'accélération V est appliquée entre la cathode chaude K et la grille cylindrique S qui l'entoure. Autour de la grille S est disposée

l'anode cylindrique A. La distance de séparation entre la cathode K et la grille S doit être petite afin que le nombre de chocs y soit petit. Par contre, l'espace entre la grille S et l'anode A doit être assez grand pour que la grande majorité des chocs s'y produisent. Pour que cet espace soit libre de tout champ électrique, on connecte l'anode à la grille. Comme dans l'appareil de Hertz (fig. 25), l'espace où les électrons sont accélérés est séparé de l'espace où ils entrent en collision avec les atomes et les excitent.

Supposons, par exemple, que le tube de travail soit rempli de vapeur de sodium. L'expérience montre que la vapeur de sodium ne devient lumineuse que si la tension d'accélération atteint 2,1 V. A cette tension, on n'arrive à exciter que la raie jaune D du sodium $\lambda = 589{,}6$ nm. Il s'ensuit que c'est la raie de résonance et que la tension $V = 2{,}1$ V est le potentiel de résonance. Pour préciser la valeur de ce potentiel on peut substituer la valeur de λ dans la formule (14.1). On trouve alors pour le potentiel de résonance du sodium $V = 2{,}103$ V. En faisant croître la tension d'accélération, on fait apparaître les autres raies du spectre d'émission du sodium.

§ 15. Emission par résonance et luminescence

1. Pour compléter les données expérimentales du paragraphe précédent, examinons encore les expériences de Wood (1868-1955), commencées en 1904-1905, sur l'existence de la luminosité dans les gaz. Wood plaçait un petit morceau de sodium métallique dans un ballon en verre où il faisait le vide. Après chauffage, le ballon était rempli de vapeur de sodium. A l'aide d'une lentille de condensateur on focalisait sur la vapeur de sodium la lumière produite par un brûleur, dans la flamme duquel on avait introduit du sel de cuisine. Cette source lumineuse émet les deux raies jaunes D_1 et D_2 du sodium, dont les longueurs d'onde sont $\lambda_{D_1} = 589{,}6$ nm, $\lambda_{D_2} = 589{,}0$ nm. Lorsque la vapeur de sodium était éclairée par ce faisceau lumineux, elle commençait à émettre une lumière jaune dans toutes les directions. L'analyse spectroscopique montra que cette lumière était composée des deux mêmes raies jaunes du sodium D_1 et D_2. On constata aussi que les raies D_1 et D_2 étaient fortement absorbées par la vapeur de sodium ; l'absorption augmentait avec la température, donc avec la densité de la vapeur de sodium, tandis que la profondeur de pénétration de la lumière incidente dans la vapeur diminuait. Dans ces conditions, la dimension longitudinale de la zone lumineuse diminuait. Lorsque la densité de la vapeur devenait grande, la zone lumineuse se réduisait à une mince couche superficielle délimitée par le faisceau incident. Simultanément, les raies D_1 et D_2 s'élargissaient puis se confondaient.

Wood observa un phénomène analogue avec la vapeur de mercure. Dans ce cas, il utilisa comme source lumineuse la radiation ultra-

violette du mercure, d'une longueur d'onde $\lambda = 253,7$ nm. Il est bien entendu que pour éviter une forte absorption de la lumière incidente par les parois du ballon contenant la vapeur de mercure, il faut que le ballon soit fabriqué en quartz et la source lumineuse soit une lampe de quartz à mercure. Wood constata que la vapeur de mercure absorbe fortement les radiations d'une longueur d'onde $\lambda = 253,7$ nm et les réémet dans toutes les directions avec la même longueur d'onde. Wood réussit à observer le même phénomène en éclairant la vapeur de mercure avec les radiations de la deuxième raie du mercure $\lambda = 185,0$ nm; l'expérience est dans ce cas beaucoup plus délicate, vu que cette raie est beaucoup plus fortement absorbée. Par la suite, des observations analogues ont pu être faites avec divers autres éléments chimiques. Il ne fait aucun doute que ce phénomène doit s'observer avec toutes les substances, quoique sa réalisation soit rendue difficile par une forte absorption et la nécessité de disposer d'une source lumineuse adéquate.

2. A l'époque où Wood procédait à ses expériences, leur interprétation ne pouvait être que classique. Tout atome possède des fréquences propres et les raies spectrales qu'il peut émettre correspondent à ces fréquences. Lorsqu'on éclaire un atome avec une lumière monochromatique d'une fréquence déterminée, des oscillations forcées y sont excitées et l'atome se met à réémettre une lumière de même fréquence. Cet effet doit être particulièrement intense lorsque la fréquence du rayonnement incident coïncide avec l'une des fréquences propres de l'atome, i.e. lorsqu'il y a résonance. C'est pour cela que ce phénomène fut appelé *émission par résonance*. On conçoit qu'un atome excité peut non seulement réémettre l'énergie qu'il a reçue sous forme de rayonnement, mais tout aussi bien la transmettre aux atomes environnants sans émettre de radiations. Cette partie de l'énergie est finalement dissipée sous forme de chaleur; cette perte d'énergie donne lieu à un affaiblissement ou à une extinction de l'émission par résonance. Cet effet est d'autant plus marqué que les interactions entre les atomes du gaz considéré sont fortes. C'est ce qui se produit lorsque la densité du gaz augmente ou lorsqu'on ajoute au gaz des atomes étrangers. Par exemple, on observe aisément l'émission par résonance de la vapeur de mercure à une pression de 0,001 mm Hg. Mais dès qu'on y ajoute de l'hydrogène sous une pression de 0,2 mm Hg, l'intensité de l'émission lumineuse diminue de deux fois et s'affaiblit encore davantage lorsque la teneur en hydrogène continue à augmenter. L'addition d'autres gaz exerce un effet semblable.

3. A première vue, l'interprétation classique de l'émission par résonance, qui vient d'être exposée, paraît raisonnable et la seule concevable. Mais un examen critique montre aussitôt les insuffisances de cette interprétation. L'émission par résonance n'est pas un phénomène indépendant, c'est un cas particulier d'un phénomène plus gé-

néral — l'*émission des raies spectrales*. Ces deux phénomènes doivent donc avoir le *même* mécanisme. La physique classique n'arriva pas à expliquer l'origine des raies et des séries spectrales, mais la théorie quantique le fit. Par conséquent, l'émission par résonance doit avoir aussi une explication quantique qui est la suivante.

Un quantum de lumière $h\nu$ tombant sur un atome doit l'exciter, i.e. le faire passer de son niveau énergétique normal $\mathscr{E}_1$ sur un niveau situé plus haut dans l'échelle de l'énergie. Soit $\mathscr{E}_2$ le niveau supérieur le plus proche du niveau $\mathscr{E}_1$. Si $h\nu < \mathscr{E}_2 - \mathscr{E}_1$, l'énergie du quantum de lumière ne peut suffire pour faire passer l'atome sur le niveau $\mathscr{E}_2$, et l'atome ne sera pas excité. Pour exciter le niveau $\mathscr{E}_2$ il faut que $\mathscr{E}_2 \leqslant \mathscr{E}_1 + h\nu$. L'excitation maximale sera réalisée pour $\mathscr{E}_2 = \mathscr{E}_1 + h\nu$, lorsque la totalité de l'énergie $h\nu$ du quantum incident est utilisée pour exciter le niveau $\mathscr{E}_2$. Au bout d'un intervalle de temps court, l'atome retourne du niveau $\mathscr{E}_2$ sur le niveau inférieur le plus proche, i.e. sur $\mathscr{E}_1$. Ce passage $\mathscr{E}_2 \rightarrow \mathscr{E}_1$ s'accompagne de l'émission d'un quantum de lumière $h\nu' = \mathscr{E}_2 - \mathscr{E}_1$. Par conséquent $\nu = \nu'$, ce qui signifie que les fréquences de la lumière excitatrice et de la lumière réémise sont égales. C'est l'explication quantique du phénomène d'émission par résonance.

Le fait que dans le cas de la vapeur de sodium on observe l'émission non pas d'une seule mais de deux raies D_1 et D_2 très rapprochées, s'explique par ce que le niveau excité $\mathscr{E}_2$ de l'atome de sodium se compose de deux niveaux énergétiques très rapprochés.

Les approches quantique et classique permettent d'expliquer l'*extinction de l'émission lumineuse* qui se produit lorsque la concentration du gaz augmente. Un atome se trouvant sur un niveau excité $\mathscr{E}_2$ peut transmettre son énergie aux atomes avoisinants sans émission de radiations. Cet effet se trouve renforcé lorsque la densité du gaz augmente ou lorsqu'on l'additionne de substances étrangères.

L'explication qui vient d'être donnée ne constitue qu'un schéma général du phénomène et n'embrasse pas certains aspects. Par exemple, on n'a pas envisagé la question de l'élargissement des niveaux énergétiques et des raies spectrales, mais c'est un sujet qui est en dehors du cadre de ce cours.

4. Les postulats de Bohr constituent la base qui permet d'expliquer les phénomènes de *luminescence* qui, auparavant, ne trouvaient aucune explication. Mais il faut d'abord préciser ce qu'est la luminescence.

Tous les corps émettent des ondes électromagnétiques dont l'intensité et la polarisation dans différentes régions du spectre dépendent de la température du corps et de son pouvoir absorbant, conformément à la loi de Kirchhoff (1824-1887). C'est le rayonnement *thermique* ou *d'équilibre*. Outre le rayonnement thermique, nombre de corps peuvent émettre un rayonnement par un mode d'excitation autre que l'excitation thermique. Citons, par exemple, la lueur de l'écran des récepteurs de télévision, la lueur du gaz contenu dans un tube à décharge lorsqu'on y fait passer un courant électrique, la lueur émise par la brisure d'un morceau de sucre, la lueur émise par certains organismes vivants, la lueur émise lors de la pourriture du bois, etc. Ce sont des exemples de *lumière*

froide. La luminescence est l'émission d'un rayonnement excédentaire par rapport au rayonnement thermique ayant une durée d'émission, après cessation de l'action extérieure qui la détermine, longue par rapport à la période des vibrations lumineuses.

La première partie de cette définition, ainsi que le terme même de luminescence, ont été avancés par G. Wiedemann (1826-1899). La deuxième partie de cette même définition est due à S.I. Vavilov, qui la formula afin de pouvoir distinguer la luminescence des phénomènes d'émission secondaire tels que réflexion et diffusion de la lumière, émission par freinage des particules chargées, émission de Vavilov-Cerenkov, qui tous ont une durée plus courte. Il s'agit là d'un critère *approximatif* qui, dans certains cas, ne permet pas de distinguer la luminescence d'autres phénomènes. Autrefois, on subdivisait la luminescence en *fluorescence* et *phosphorescence*. On désignait par fluorescence les rayonnements qui cessent en même temps que l'action de l'excitation extérieure. Lorsque la luminosité subsistait après la fin de l'excitation, le phénomène était appelé phosphorescence. On s'est aperçu par la suite qu'il n'y avait aucune différence de principe entre la fluorescence et la phosphorescence, la distinction étant purement quantitative ne se rapportant qu'à la durée de postluminescence. Cette division en deux catégories est abandonnée aujourd'hui. Le terme « fluorescence » est encore utilisé, mais uniquement pour faire ressortir la brièveté de la luminescence. Par exemple, la durée de l'émission lumineuse par résonance des gaz raréfiés est de l'ordre de 10^{-8} à 10^{-9} s.

La définition que nous venons de donner de la luminescence appelle quelques remarques. Comme la température est un *concept macroscopique*, on ne doit diviser le rayonnement total en émission thermique et en luminescence que lorsqu'on a en vue non pas des molécules et des atomes isolés, mais des *ensembles* de molécules et d'atomes, donc des corps dont on peut déterminer la température. Si le corps se trouve dans un état hors d'équilibre, tel qu'il ne peut être question de sa température, on ne peut parler ni de rayonnement thermique ni de luminescence. On ne pourra en parler non plus dans le cas d'émission lumineuse par des atomes ou des molécules isolés. C'est pour cela que, parlant de la luminosité des gaz, qui a été l'objet des recherches de Wood, nous avons préféré recourir au terme neutre d'« émission par résonance » au lieu de « fluorescence par résonance » que l'on utilise parfois. En effet, on observe une telle émission lumineuse dans les gaz raréfiés, tandis que l'émission dans les gaz relativement denses pourrait être appelée « fluorescence par résonance ».

Les substances où on arrive à exciter la luminescence sont appelées substances luminescentes ou *luminophores*. Selon le mode d'excitation, on distingue:

1) la *photoluminescence* qui est la luminescence qui se manifeste lorsqu'on éclaire une substance luminescente avec de la lumière visible ou ultraviolette;

2) la *luminescence* X qui est la luminescence qui se manifeste lorsqu'une substance luminescente est irradiée par des rayons X;

3) la *luminescence cathodique* est provoquée par le bombardement d'une substance luminescente par des électrons (rayons cathodiques);

4) la *radioluminescence* résulte de l'irradiation par des radiations nucléaires: particules α, particules β, rayons gamma, protons, neutrons, etc.;

5) l'*électroluminescence*, dont un exemple est la luminescence d'un tube à décharge à gaz sous l'action d'un champ électrique;

6) la *chimiluminescence* qui accompagne certaines réactions chimiques;

7) la *triboluminescence* qui concerne l'émission de lumière qui accompagne des actions mécaniques, tels le broyage, l'écrasement et le clivage de certains cristaux (de sucre, par exemple);

8) la *candoluminescence* qui est une émission lumineuse (non thermique) qui se manifeste lorsqu'on place certaine substance dans une flamme (d'un brûleur à gaz par exemple).

5. Il n'est pas opportun de justifier ici la classification ci-dessus, car la luminescence en tant que phénomène général, ainsi que certains types particuliers de luminescence, font l'objet de nombreux ouvrages qu'on ne peut analyser ici.

Ce qui importe, c'est de noter le *caractère quantique* de ce phénomène. Comme tout autre rayonnement, la luminescence résulte de transitions quantiques du système émetteur entre un niveau énergétique excité et un niveau d'énergie plus petite. Cette approche permet d'expliquer l'une des principales caractéristiques de la luminescence — sa *durée*, car il existe de nombreux corps pouvant se trouver longtemps dans un état excité. Le type de luminescence est déterminé par la structure du spectre énergétique du corps, par la durée moyenne de son séjour à l'état excité, par les règles de sélection valables dans l'absorption et l'émission de lumière, etc.

En 1852, Stokes (1819-1903) énonça une règle selon laquelle *les radiations de luminescence ont une longueur d'onde plus grande que les radiations lumineuses excitatrices*: $\lambda' > \lambda$. Cette règle témoigne du caractère quantique de la photoluminescence. En effet, si un quantum de lumière excitatrice fait passer le système du niveau normal $\mathscr{E}_1$ à un niveau excité $\mathscr{E}_2$, on a $\mathscr{E}_2 \leqslant \mathscr{E}_1 + h\nu$. Lors du retour du système dans l'état initial $\mathscr{E}_1$, la fréquence de la lumière émise sera $\nu' = (\mathscr{E}_2 - \mathscr{E}_1)/h$, et par suite $\nu' \leqslant \nu$. Au retour de l'état excité le système, au lieu de revenir dans l'état initial, peut tomber dans un niveau se trouvant un peu au-dessus du niveau $\mathscr{E}_1$. On aura alors $\nu' < \nu$.

On observe cependant des écarts à la règle de Stokes. Supposons qu'un quantum de lumière fait passer le système d'un niveau excité $\mathscr{E}'_1$ sur un niveau $\mathscr{E}'_2$ situé plus haut dans l'échelle des énergies. On aura alors $\mathscr{E}'_2 \leqslant \mathscr{E}'_1 + h\nu$. Mais comme $\mathscr{E}'_1 > \mathscr{E}_1$, il peut arriver que $\mathscr{E}'_2 > \mathscr{E}_1 + h\nu$, et le retour du système sur le niveau initial $\mathscr{E}_1$ s'accompagnera de l'émission d'un quantum de fréquence $\nu' > \nu$.

§ 16. Insuffisances de principe de la théorie de Bohr

La théorie de Bohr constitua une étape importante dans la compréhension des phénomènes intra-atomiques. La physique classique, qui fut élaborée sur la base de phénomènes macroscopiques, se heurta à des difficultés de principe dès qu'on chercha à l'appliquer aux atomes, aux molécules et d'une façon générale aux phénomènes du microcosme. La théorie de Bohr prépara les esprits à accepter l'idée que *les conceptions et les lois classiques ne peuvent suffire à la compréhension des phénomènes du microcosme*. L'étude du microcosme réclamait de *nouveaux concepts* et de *nouvelles lois*. Un rôle de premier plan devait revenir au quantum d'action découvert par Planck. La théorie de Bohr stimula grandement le développement des recherches expérimentales qui apportèrent une riche moisson de résultats. Même dans les cas — et ils étaient nombreux — où la théorie n'était pas en mesure de donner des interprétations quantitatives des phénomènes observés, les deux postulats de Bohr servaient de guide pour la classification et l'interprétation qualitative des phénomènes. C'est ainsi que les postulats de Bohr servirent de base pour la classification d'une énorme masse de données empiriques concernant la spectroscopie atomique et moléculaire.

Il est certain que les deux postulats de Bohr ne pouvaient suffire à bâtir une théorie complète. Il a fallu les compléter tout d'abord par des *règles de quantification* à l'aide desquelles il devenait possible de calculer les niveaux énergétiques de l'atome. Bohr avança une règle de quantification des orbites électroniques cir-

culaires pour les atomes monoélectroniques; cette règle s'exprime par la formule (13.6). Bohr reconnaissait lui-même que lorsqu'il cherchait cette formule, il était dominé par le désir d'adapter la théorie à la formule empirique de Balmer pour les termes spectraux de l'hydrogène. Peu de temps après, Sommerfeld généralisa la règle de quantification de Bohr au mouvement des électrons sur des orbites elliptiques, mais cette nouvelle règle n'était applicable qu'aux atomes à un seul électron. On n'arrivait pas à étendre la règle de quantification aux atomes à plusieurs électrons, même à l'atome d'hélium, le plus simple après l'atome d'hydrogène, comportant un noyau et deux électrons. On ne peut évidemment exiger que la théorie donne une solution analytique au problème des trois corps, car une telle solution peut être irréalisable. Mais la théorie doit fournir une *méthode de principe* pour trouver une solution numérique avec une précision suffisante pour être confrontée aux résultats des expériences *). La théorie de Bohr n'arrivait pas à le faire. Encore moins pouvait-elle donner une solution de principe au problème, plus compliqué, de la formation d'une molécule aussi simple que la molécule d'hydrogène formée de deux protons et de deux électrons.

Même dans le cas simple d'atomes à un seul électron, la théorie de Bohr ne permettait de calculer que les fréquences des raies spectrales sans pouvoir calculer les intensités et la polarisation de ces raies. Il est vrai que pour calculer les intensités et la polarisation, on faisait appel au *principe de correspondance*. Mais comme ce principe n'est justifié que pour les grands nombres quantiques, le calcul pouvait être effectué par les méthodes classiques. Néanmoins la théorie de Bohr étendait ce principe aux petits nombres quantiques, ce qui était dénué de tout fondement. On était ainsi amené à déterminer les intensités et la polarisation par les méthodes classiques.

Le défaut majeur de la théorie de Bohr est qu'elle est *inconséquente*. Elle ne reconnaissait que l'existence des états stationnaires de l'atome, ou comme le disait Bohr, d'orbites stationnaires des électrons. C'est une proposition absolument étrangère à la mécanique classique, et néanmoins la théorie décrivait le mouvement des électrons à l'état stationnaire à l'aide des lois de la mécani-

---

*) Tel est le cas en mécanique céleste, où on n'arriva pas à trouver une solution analytique au problème des trois corps. En mécanique céleste, on élabora des méthodes approchées efficaces de solution numérique du problème, dont la précision est suffisante. Notons qu'en 1912, Sundman trouva une solution analytique sous forme de séries. Mais ces séries ne convenaient pas au calcul numérique parce qu'elles ne convergeaient que très lentement. Pour que ces séries fournissent la même précision que celle que donnent aujourd'hui les annuaires astronomiques, on devrait calculer la somme de $10^{8\,000\,000}$ termes. On ne peut faire ce calcul même avec les meilleures calculatrices électroniques modernes.

que classique, tout en rejetant l'électrodynamique classique (pour la raison que le rayonnement est absent). Selon une remarque ironique de H. Bragg (1862-1942), la théorie de Bohr utilise les lois classiques tous les lundis, mercredis et vendredis, et la mécanique quantique les autres jours. Si on rejette le concept d'orbites électroniques dans les atomes — et c'est ce qui a été fait plus haut — ces deux postulats sont *confirmés par l'expérience* et doivent donc être considérés comme *justes*. Dans son ensemble, la théorie de Bohr n'est qu'une étape intermédiaire sur la voie menant à une théorie plus parfaite et plus conséquente. Bohr le comprenait mieux que quiconque et c'est lui qui a eu le mérite de chercher la vraie signification des fondements de la mécanique quantique qui vint se substituer à sa théorie.

C'est Planck qui introduisit l'idée du caractère quantique des processus d'émission et d'absorption de la lumière. Einstein étendit la notion de quantification de la lumière à sa propagation dans l'espace en avançant le concept de photon. La théorie de la capacité calorifique d'Einstein démontra directement que la constante h introduite par Planck apparaissait non seulement dans les phénomènes lumineux, mais aussi dans les processus évoluant dans les substances. Un nouveau pas en avant fut fait par Bohr. Les succès que remporta la théorie de Bohr dans le calcul de la constante de Rydberg et dans celui des dimensions de l'atome révélèrent que la constante de Planck devait être considérée comme une *grandeur fondamentale universelle* figurant dans la description de toutes les formes de la matière, et non seulement dans l'interprétation du dualisme de la lumière. En effet, la constante de Planck ayant des dimensions de l'action (dimensions du produit de l'impulsion par la coordonnée, ou du produit de l'énergie par le temps) figure dans toutes les parties de la physique moderne.

CHAPITRE III

PROPRIÉTÉS ONDULATOIRES DES PARTICULES DE MATIÈRE

§ 17. L'hypothèse de L. de Broglie

1. L'élaboration d'une théorie conséquente pour la description des phénomènes à l'échelle atomique et subatomique a été réalisée en gros dans les années 1925-1926. Cette théorie fut appelée *mécanique quantique*. Mais d'abord fut formulée la *mécanique des matrices*, qui n'est qu'une des voies de développement des conceptions quantiques. Les idées de la mécanique des matrices ont été exposées dans un écrit fondamental de Heisenberg (1901-1976) « Sur l'interprétation quantomécanique des relations cinématiques et mécaniques ». Une construction systématique de la mécanique des matrices fut réalisée par Born (1882-1970) et Jordan (1902-1980) que rejoignit plus tard Heisenberg. A cette voie de développement de la mécanique quantique s'apparente la théorie qui fut élaborée simultanément, mais de façon indépendante, par Dirac (1902-1984). Quelques années plus tard, une autre voie de développement de la théorie fut initiée par Schrödinger (1887-1961), que l'on appela *mécanique ondulatoire*. On s'aperçut bientôt que ces deux voies de développement, différentes par la forme, étaient identiques du point de vue contenu physique. Dans un cours de physique générale, il n'est pas possible d'exposer les idées très abstraites de la mécanique des matrices, mais il est toujours utile de donner un exposé, même bref, des conceptions physiques de la mécanique ondulatoire. Bien entendu, on ne peut exposer les détails de l'appareil mathématique compliqué de cette théorie, qui constitue une partie inhérente et essentielle de la mécanique quantique : cela est du ressort du cours de physique théorique.

2. La construction par Schrödinger de la mécanique ondulatoire a eu pour point de départ les travaux de L. de Broglie (né en 1892). En 1923-1924, L. de Broglie avança et développa l'idée des ondes de matière. En ce temps, l'optique se trouvait confrontée à une situation paradoxale, mais étayée par des données factuelles. Dans certains phénomènes (interférences, diffraction, etc.) la lumière se comporte comme les *ondes*, tandis que d'autres phénomènes (effet photoélectrique, effet Compton, etc.) témoignent de façon aussi certaine que la lumière se comporte comme si elle était composée de *particules*. De Broglie se demanda si ce dualisme onde-corpuscule pouvait se manifester dans le cas de la matière ordinai-

re et s'il en était bien ainsi, quelles seraient les propriétés ondulatoires des particules de matière. Les expériences ultérieures répondirent à cette question par l'affirmative.

Soit une particule se déplaçant avec la vitesse $\boldsymbol{v}$ dans l'espace libre. De Broglie supposa qu'à cette particule était associée une certaine onde monochromatique plane définie par

$$\Psi = \Psi_0 e^{i(\boldsymbol{kr}-\omega t)}, \tag{17.1}$$

qui se propagerait dans la direction de la vitesse $\boldsymbol{v}$ *). De Broglie ne pouvait préciser la nature physique de cette onde (i.e. préciser la signification physique de la fonction Ψ). On laissera aussi provisoirement de côté cette question. Les ondes de la forme (17.1) ont été appelées *ondes de phase*, *ondes de matière* ou *ondes de de Broglie.*

Essayons d'établir une relation rationnelle entre les propriétés corpusculaires et ondulatoires d'une particule, qui soit complètement indépendante de la nature physique de la grandeur Ψ. Exigeons que cette relation satisfasse à la condition d'invariance relativiste. Les propriétés corpusculaires de la particule sont caractérisées par son énergie $\mathscr{E}$ et son impulsion $\boldsymbol{p}$, et ses propriétés ondulatoires sont caractérisées par la fréquence ω et le nombre d'onde $\boldsymbol{k}$. On entendra par $\mathscr{E}$ l'*énergie totale* de la particule dans le sens qu'attribue à ce terme la théorie de la relativité. Pour que la définition soit parfaitement univoque, on doit exiger que *l'énergie et l'impulsion forment le quadrivecteur* ($\mathscr{E}/c$, $\boldsymbol{p}$) (voir t. IV, § 111,3). Pour définir la fréquence ω on exigera que la phase de l'onde $\omega t - \boldsymbol{kr}$ satisfasse à la condition d'invariance relativiste (voir à ce sujet § 19,9); dans ce cas, ω et $\boldsymbol{k}$ pourront définir le quadrivecteur (ω/c, $\boldsymbol{k}$). Si on exige encore que les composantes temporelles et spatiales des quadrivecteurs ($\mathscr{E}/c$, $\boldsymbol{p}$) et (ω/c, $\boldsymbol{k}$) soient proportionnelles les unes aux autres, on arrive à des relations relativistes invariantes

$$\mathscr{E} = \hbar\omega, \tag{17.2}$$

$$\boldsymbol{p} = \hbar\boldsymbol{k}. \tag{17.3}$$

Ces relations seront identiques aux relations correspondantes pour les photons, en posant que pour toutes les particules $\hbar$ est égal à la constante de Planck. Ce choix de la valeur de $\hbar$ n'est pas logiquement indispensable, mais se trouve justifié par les résultats auxquels il conduit. De Broglie postula les relations (17.2) et (17.3).

---

*) En optique, on définissait une onde monochromatique par une expression de la forme $e^{i(\omega t-\boldsymbol{kr})}$. Actuellement, on utilise l'expression conjuguée complexe $e^{i(\boldsymbol{kr}-\omega t)}$. Ces deux expressions sont équivalentes, mais en mécanique quantique il est d'usage d'écrire les ondes sous la forme (17.1). Indépendamment du mode d'écriture, on doit entendre par phase de l'onde l'expression $\omega t - \boldsymbol{kr}$.

Dans tout référentiel d'inertie, le vecteur d'onde k est défini univoquement puisque la relation (17.3) le relie de façon univoque à l'impulsion $p = mv$ de la particule. Par contre, la relation (17.2) ne présente pas un caractère aussi parfaitement univoque, car celui-ci est imposé de façon artificielle en exigeant que $\mathcal{E}$ et ω soient des composantes temporelles des quadrivecteurs. Or, en principe, l'énergie est toujours définie à une constante additive près. On verra plus tard (cf. § 19) qu'on peut trouver une définition de la fréquence ω telle que celle-ci contienne aussi une constante additive.

3. Examinons quelques propriétés des ondes de de Broglie découlant des relations (17.2) et (17.3). On tira tout d'abord de (17.3) une expression de la longueur des ondes de de Broglie

$$\lambda = \frac{2\pi}{k} = \frac{2\pi\hbar}{p} = \frac{h}{p}. \tag{17.4}$$

Dans chaque référentiel d'inertie, cette longueur d'onde est définie de façon univoque.

Les formules (17.2) et (17.3) donnent pour la vitesse de phase de l'onde de de Broglie la relation suivante:

$$v_{\text{ph}} = \frac{\omega}{k} = \frac{\mathcal{E}}{p}. \tag{17.5}$$

En théorie relativiste, $\mathcal{E} = mc^2$, $p = mv$, où v est la vitesse de la particule et m sa masse relativiste. On aura donc

$$v_{\text{ph}} = \frac{c^2}{v}. \tag{17.6}$$

Puisqu'on a toujours $v \leqslant c$, on doit avoir $v_{\text{ph}} \geqslant c$. Pour les photons dans le vide, $v = c$ et par suite $v_{\text{ph}} = c$. Ce dernier résultat ne doit pas étonner puisqu'aucune contrainte n'est imposée à la vitesse de phase. D'ailleurs, on montrera par la suite que selon l'interprétation moderne, la vitesse de phase des ondes de de Broglie ne présente qu'une signification toute symbolique, pour la raison que, selon cette interprétation, la vitesse de phase est, par principe, une grandeur non observable.

La grandeur en principe observable est la vitesse de groupe des ondes de de Broglie

$$v_{\text{gr}} = \frac{d\omega}{dk} = \frac{d\mathcal{E}}{dp}. \tag{17.7}$$

Cette grandeur ne comporte aucune indétermination, puisque aussi bien dp que les accroissements de l'énergie $d\mathcal{E}$ sont définis univoquement. Quelle que soit la vitesse de la particule $d\mathcal{E} = v\,dp$, de sorte qu'on a toujours

$$v_{\text{gr}} = v, \tag{17.8}$$

ce qui signifie que la vitesse de groupe des ondes de de Broglie est égale à la vitesse de la particule. En remplaçant v par v_{gr} dans la formule (17.6), on obtient

$$v_{ph} \cdot v_{gr} = c^2. \qquad (17.9)$$

Tout formellement, on peut définir une grandeur analogue à la longueur de l'onde de de Broglie (17.4). Pour cela il suffit de remarquer que dans l'univers de Minkowski la longueur du quadrivecteur énergie-impulsion de la particule est égale à $\sqrt{(\mathcal{E}/c)^2 - p^2}$. C'est un invariant égal à m_0c, où m_0 est la masse au repos de la particule. En divisant la constante de Planck h par cet invariant, on obtient la quantité invariante

$$\lambda_C = \frac{h}{m_0c}, \qquad (17.10)$$

qui a la dimension de la longueur. Cette quantité représente la longueur de Compton de la particule. Tout formellement, on peut la considérer comme la longueur de l'onde de de Broglie correspondant à une impulsion égale à la longueur invariante du quadrivecteur énergie-impulsion de la particule dans l'univers de Minkowski.

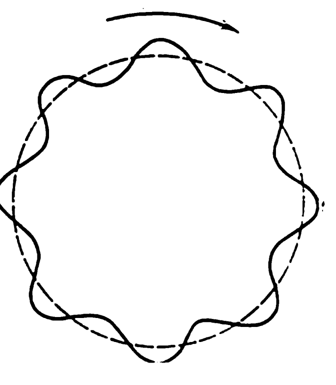


Fig. 27

4. De Broglie utilisa le concept d'ondes de phase pour concrétiser le contenu de la règle de quantification de Bohr (13.6) appliquée à un atome à un seul électron, qui paraissait jusqu'alors quelque peu mystérieuse. Il considéra une onde de phase progressant autour du noyau sur l'orbite circulaire de l'électron. Si la longueur d'onde λ est un multiple entier de la longueur de la circonférence (fig. 27), à chaque tour l'onde revient au point de départ avec la même phase et la même amplitude. Un régime vibratoire invariable dans le temps s'établira alors à chaque point de l'orbite, le rayonnement sera nul et l'orbite sera stationnaire. Si cette condition n'est pas satisfaite, à la suite d'un parcours circulaire complet autour du noyau la phase et l'amplitude de l'onde ne retrouveront pas leurs valeurs initiales et l'état ne sera plus stationnaire. Partant de ce raisonnement, de Broglie exprima la règle de quantification, imposant le caractère stationnaire de l'orbite, sous la forme

$$\frac{2\pi R}{\lambda} = n. \qquad (17.11)$$

où R est le rayon de l'orbite circulaire et n un nombre entier (nombre quantique principal). En y posant $\lambda = h/p = 2\pi\hbar/p$ et en remar-

quant que $L = Rp$ est le moment cinétique de l'électron, on obtient

$$L = n\hbar, \tag{17.12}$$

résultat identique à (13.6). Pour de Broglie, cette concordance témoignait de la justesse de son concept d'onde de phase. Par la suite, on réussit à généraliser la condition (17.11) au cas d'orbites elliptiques sur lesquelles la longueur d'onde λ varie le long de la trajectoire de l'électron. On était en droit de penser que le succès de la théorie en était renforcé.

En fait ce succès était illusoire. Le raisonnement de de Broglie présuppose que l'onde se propage non pas dans l'espace, mais le long d'une ligne représentant l'orbite stationnaire de l'électron. Une telle idéalisation correspond à l'approximation de l'optique géométrique, que l'on peut utiliser dans le cas limite où la longueur d'onde λ est négligeable devant le rayon de l'orbite de l'électron, i.e. pour de grandes valeurs du nombre quantique. Or, dans ce dernier cas, le problème de la quantification ne présente plus d'intérêt. Pour arriver à un résultat essentiellement nouveau, il faut remplacer l'optique géométrique par l'optique ondulatoire. C'est ce que fit Schrödinger.

5. On peut arriver aux mêmes résultats en suivant une autre voie. Introduisons à cet effet l'*indice de réfraction* μ des ondes de de Broglie, qui est une quantité importante présentant de l'intérêt par elle-même. Convenons d'appeler *milieu* l'espace dans lequel se propage l'onde de de Broglie. Si le milieu ne contient aucun champ de force, il sera homogène. L'indice de réfraction du milieu ne peut être défini qu'à un facteur constant près, puisque seul importe le rapport des indices de réfraction des milieux à la frontière desquels se produit la réfraction des ondes de phase. Dans toutes les théories ondulatoires μ est inversement proportionnel à la vitesse de phase de l'onde. Pour les ondes de de Broglie, $\mu \sim 1/v_{ph} = v/c^2$. On posera, en omettant un facteur constant, que

$$\mu = v. \tag{17.13}$$

Convenons d'appeler *absolu* l'indice de réfraction défini par cette formule. La formule (17.13) conserve toute sa signification même dans le cas où la vitesse v de la particule change d'un point à l'autre, donc en présence de champs de force. La vitesse v, et donc μ, sont définis univoquement en chaque point par l'équation de l'énergie $\mathcal{E} + U =$ const, qui présuppose que la fonction potentielle U ne dépend que des coordonnées, mais ne dépend pas de façon explicite du temps.

Dans le cas limite de courtes longueurs d'onde, la propagation des ondes se produit le long de lignes ou *rayons* indépendants (voir t. IV, § 6). Ce cas correspond à l'*optique géométrique*. La propagation d'une perturbation ondulatoire le long des rayons est formellement

analogue au mouvement d'une particule de la mécanique classique le long d'une trajectoire. Le rayon de courbure R du rayon ou de la trajectoire d'une particule est donné par la formule

$$\frac{1}{R}=\frac{\partial}{\partial N}(\ln \mu)=\frac{\partial}{\partial N}(\ln v). \qquad (17.14)$$

On doit différentier cette expression suivant la direction de la normale principale N au rayon ou à la trajectoire (voir t. IV, § 4). On peut utiliser cette formule à la place des équations de Newton pour déterminer la forme du rayon ou de la trajectoire de la particule.

Revenons à l'établissement de la règle de quantification de de Broglie. La condition d'applicabilité de l'optique géométrique au mouvement de l'électron autour du noyau atomique s'exprime par la formule

$$\left|\lambda \frac{d\mu}{dr}\right| \ll \mu, \quad \text{i.e.} \quad \left|\lambda \frac{dv}{dr}\right| \ll v.$$

En y substituant $\lambda = 2\pi\hbar/p$, $\mu = v$ et en se limitant à l'approximation non relativiste, on écrira

$$\left|2\pi\hbar \frac{dv}{dr}\right| \ll pv = 2K,$$

où K est l'énergie cinétique de l'électron. On trouve la vitesse v à partir de l'équation de l'énergie

$$\frac{mv^2}{2}-\frac{Ze^2}{r}=\mathscr{E}=\text{const.}$$

On en déduit

$$mvr\frac{dv}{dr}=-\frac{Ze^2}{r}=U,$$

où U est l'énergie potentielle. Si l'électron se meut sur une orbite circulaire, $mvr = n\hbar$, et par suite

$$n\hbar\frac{dv}{dr}=U=\mathscr{E}-K=-2K,$$

car $\mathscr{E} + K = 0$ pour un mouvement circulaire. En identifiant cette relation avec la dernière inégalité ci-dessus, on trouve

$$n \gg 2\pi,$$

ce qui signifie que le nombre quantique n doit être grand; ce résultat confirme l'assertion qui a été énoncée plus haut.

6. Ce qui précède n'est qu'une vue de l'esprit, une construction hypothétique et ne possède de ce fait aucun pouvoir probant. La preuve ou la réfutation des résultats théoriques obtenus ne peut résulter que d'expériences. Il faut donc choisir convenablement les phénomènes où les propriétés ondulatoires de la matière, au cas où elles

existent bien, pourraient se manifester. Indépendamment de la nature physique des ondes, les phénomènes d'interférences et de diffraction devraient convenir, car la grandeur directement observable est la longueur d'onde λ. Dans tous les cas, la longueur de l'onde de de Broglie est définie par la formule (17.4). Appliquons cette formule au mouvement non relativiste des particules.

Pour les électrons accélérés par une différence de potentiel V, l'impulsion est $p = \sqrt{2m_e eV}$, de sorte que dans le système absolu des unités de mesure

$$\lambda_e = \frac{h}{\sqrt{2m_e eV}}. \qquad (17.15)$$

En posant $hc = 1{,}2399 \cdot 10^{-4}$ eV·cm, $m_e c^2 = 511\,003$ eV, on obtient une formule directement utilisable

$$\lambda_e = \sqrt{\frac{150{,}42}{V_{\text{volt}}}} \cdot 10^{-8}\ \text{cm} = \frac{1{,}2264}{\sqrt{V_{\text{volt}}}}\ \text{nm}. \qquad (17.16)$$

Pour les protons, on trouve de même

$$\lambda_p = \frac{0{,}02862}{\sqrt{V_{\text{volt}}}}\ \text{nm}. \qquad (17.17)$$

Calculons encore la longueur de l'onde de de Broglie pour les molécules d'un gaz au repos é la tempàrature absolue T. Ce problème n'est qu'imparfaitement défini puisque les molécules se meuvent avec des vitesses thermiques correspondant à la loi de distribution de Maxwell (1831-1879). Nous entendrons, sans le justifier ici (voir problème à la fin du paragraphe suivant), par v la vitesse quadratique moyenne des molécules. L'impulsion d'une molécule sera alors $p = \sqrt{3mkT}$. On trouve facilement pour les atomes d'hélium $m_{He} = 6{,}7 \cdot 10^{-24}$ g)

$$\lambda_{He} = \frac{1{,}26}{\sqrt{T}}\ \text{nm}, \qquad (17.18)$$

pour les molécules d'hydrogène

$$\lambda_{H_2} = \frac{1{,}78}{\sqrt{T}}\ \text{nm} \qquad (17.19)$$

et pour les neutrons thermiques

$$\lambda_n = \frac{2{,}52}{\sqrt{T}}\ \text{nm}. \qquad (17.20)$$

Ces formules montrent que pour les électrons accélérés par une différence de potentiel de 100 à 10 000 V, pour les atomes d'hélium et pour les molécules d'hydrogène à la température ordinaire, ainsi que pour les neutrons thermiques et d'autres particules légères « lentes », la longueur des ondes de de Broglie est de l'ordre de la longueur d'on-

de de rayons X mous. Il convient donc de chercher à déceler la diffraction de ces particules par les mêmes méthodes que celles mises en œuvre pour l'étude de la diffraction des rayons X. L'hypothèse de de Broglie paraissait à tel point fantastique que pendant un temps relativement long aucun expérimentateur ne se hasarda à la soumettre à une vérification expérimentale.

PROBLÈMES

1. Généraliser les formules non relativistes (17.15), (17.16) et (17.17) au cas d'électrons et de protons relativistes. Pour quelles valeurs de la différence de potentiel d'accélération V peut-on utiliser les formules non relativistes sans que l'erreur soit supérieure à 1 % ?

Réponse

$$\lambda = \frac{h}{\sqrt{2m_0eV}}\left(1-\frac{eV}{2m_0c^2}\right)^{-1/2}, \qquad (17.15a)$$

où m_0 est la masse au repos de la particule. Pour les électrons

$$\lambda_e = \frac{1,2264}{\sqrt{V_{\text{volt}}}}(1+0,978\cdot10^{-6}V_{\text{volt}})^{-1/2}\ \text{nm}, \qquad (17,16a)$$

et pour les protons

$$\lambda_p = \frac{0,02862}{\sqrt{V_{\text{volt}}}}(1+0,533\cdot10^{-9}V_{\text{volt}})^{-1/2}\ \text{nm}. \qquad (17.17a)$$

A la précision imposée, on peut utiliser les formules non relativistes lorsque $V \leqslant 20$ keV pour les électrons et $V \leqslant 40$ MeV pour les protons.

2. Trouver une expression approchée de la longueur de l'onde de de Broglie d'une particule ultrarelativiste, i.e. d'une particule dont l'énergie cinétique $\mathscr{E}_{\text{cin}}$ est beaucoup plus grande que l'énergie au repos m_0c^2. Pour quelles valeurs de $\mathscr{E}_{\text{cin}}$ l'erreur commise en utilisant la formule obtenue ne sera-t-elle pas supérieure à 1 % ? Calculer λ pour une particule ultrarelativiste d'énergie cinétique $\mathscr{E}_{\text{cin}} = 10$ GeV.

Réponse. $\lambda = hc/\mathscr{E}_{\text{cin}}$. Pour $\mathscr{E}_{\text{cin}} > 100m_0c^2$, l'erreur ne sera pas supérieure à 1 %. Pour $\mathscr{E}_{\text{cin}} = 10$ GeV, $\lambda = 1,25\cdot10^{-14}$ cm.

3. Pour quelle vitesse d'une particule les longueurs d'onde de Compton et de de Broglie seront-elles égales?

Réponse. $v = c/\sqrt{2}$.

§ 18. Confirmation expérimentale de l'hypothèse de L. de Broglie

1. Les interférences des électrons réfléchis par des cristaux ont été décelées avant que fut énoncée l'hypothèse de de Broglie, mais on ne sut pas les expliquer. En 1921-1923, Davisson (1881-1958) et Kunsman, qui effectuaient des expériences sur la diffusion des électrons par des feuilles métalliques minces, avaient constaté l'existence d'une nette dépendance de l'intensité du faisceau diffusé avec l'angle de diffusion. Les positions et les hauteurs des maximums apparaissant sur la courbe de diffusion dépendaient largement de la vitesse des électrons. Au cours d'une expérience où les électrons

étaient diffusés par une plaque de nickel, le ballon de verre éclata et la plaque de nickel s'oxyda. Au cours d'un long recuit sous vide et en atmosphère d'hydrogène, la plaque de nickel subit une recristallisation avec formation de plusieurs gros cristaux. Lorsqu'on réutilisa cette plaque pour la diffusion des électrons, on constata que la forme de la courbe de diffusion se modifia fortement — le nombre de maximums augmenta fortement et les maximums devinrent plus aigus. La figure 28 représente les diagrammes polaires de la diffusion

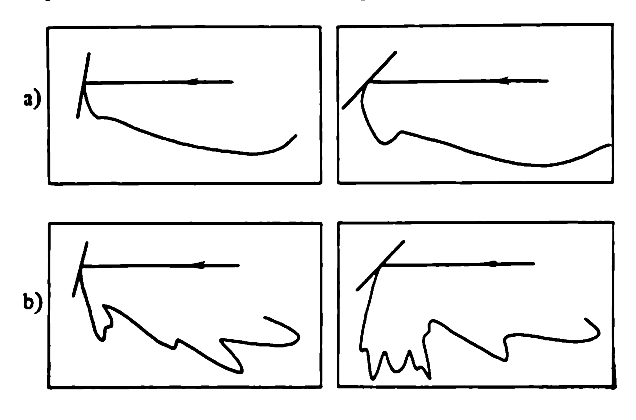


Fig. 28

des électrons : a) avant recuit de la plaque, et b) après recuit. Sur ces diagrammes on a porté le long des rayons vecteurs des segments proportionnels à l'intensité de diffusion des électrons dans les directions correspondantes. Les causes d'apparition des maximums et des minimums sur les courbes de diffusion restèrent inexpliquées jusqu'au jour où on les attribua aux interférences entre les ondes de de Broglie réfléchies par différents plans atomiques de gros cristaux formés par recristallisation lors du recuit. Cette interprétation fut confirmée, en 1927, par les expériences de Davisson et Germer (1896-1971), qui découvrirent le phénomène de diffraction des électrons. Cette découverte marqua le début d'une étude systématique de ce phénomène.

2. Dans leurs expériences, Davisson et Germer appliquèrent aux ondes de de Broglie la *méthode de L. Bragg* (1890-1971) élaborée pour les rayons X. Un faisceau parallèle d'électrons animés d'une même vitesse produit par le « canon à électrons » *A* (fig. 29) était envoyé sur le monocristal de nickel. Les électrons diffusés étaient captés par un collecteur *C* relié à un galvanomètre. Pour disposer le collecteur sous un angle quelconque par rapport à la direction du faisceau incident, on le faisait tourner dans un même plan (le plan de la figure). Les indications du galvanomètre permettaient de juger de l'intensi-

té des faisceaux électroniques diffusés dans différentes directions. Un diagramme polaire typique de l'intensité de diffusion des électrons est représenté sur la figure 30. On y discerne un maximum aigu correspondant à la réflexion spéculaire des électrons lorsque l'angle d'incidence est égal à l'angle de réflexion. En répétant la même ex-

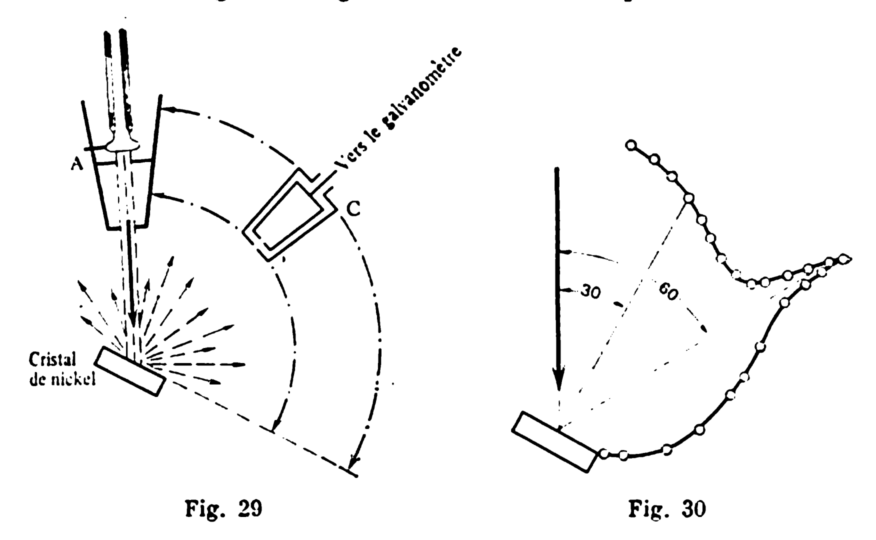


Fig. 29 Fig. 30

périence avec une plaque de nickel polycristalline composée d'une multitude de petits cristaux orientés au hasard, on ne décela aucune direction privilégiée pour la réflexion des électrons.

Les expériences de Davisson et Germer ont permis de soumettre l'hypothèse de de Broglie à une vérification quantitative. On sait (cf. t. IV, § 61) que la réflexion des rayons X par les cristaux présente un *caractère interférentiel*. Les ondes renvoyées par les différents plans atomiques parallèles du cristal sont identiques à celles que l'on obtiendrait par une réflexion spéculaire sur chacun de ces plans. Si on a la condition de Bragg-Wulff (1863-1925)

$$2d \sin \varphi = m\lambda, \tag{18.1}$$

où φ est l'angle de glissement, d la distance interréticulaire, $m = 1, 2, 3, \ldots$, ces ondes se renforcent mutuellement par interférences et donnent naissance à une onde réfléchie. Pour que ce mécanisme de réflexion soit valable, il faut que soit vérifiée l'inégalité $m\lambda/(2d) < 1$, ce qui implique la petitesse de la longueur d'onde. Cette inégalité est vérifiée pour les ondes de de Broglie si les électrons sont accélérés par une différence de potentiel de quelques dizaines ou centaines de volts. Il s'ensuit que même si on ne connaît qu'imparfaitement le mécanisme de réflexion des ondes de de Broglie, on

peut s'attendre à un phénomène d'interférences et, par voie de conséquence, à ce que la condition (18.1) s'applique aussi à ces ondes *).

Dans le cas des rayons X monochromatiques, on maintient constante la longueur d'onde λ pendant toute la durée de l'expérience et on fait varier l'angle de glissement φ jusqu'à l'apparition de la réflexion interférentielle. Avec les ondes de de Broglie il est beaucoup

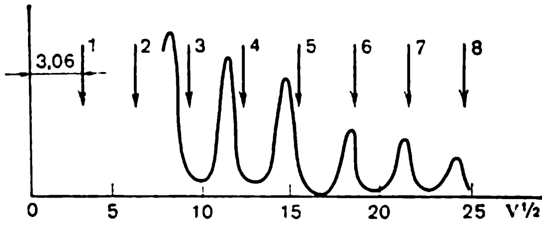


Fig. 31

plus commode de maintenir constant l'angle φ et d'obtenir une réflexion interférentielle en faisant varier la tension d'accélération, i.e. la longueur λ de l'onde de de Broglie. C'est ce que firent Davisson et Germer.

Selon la théorie, les maximums de réflexion ne doivent apparaître que pour les valeurs de λ correspondant à la formule (18.1) pour les valeurs entières de m. En substituant dans cette formule l'expression de λ tirée de (17.16), on obtient pour les électrons non relativistes

$$\sqrt{V} = \frac{1,226}{2d \sin \varphi}, \tag{18.2}$$

où V est exprimé en volts et d en nanomètres. La figure 31 représente la courbe obtenue par Davisson et Germer avec un monocristal de nickel pour $\varphi = 80°$ et $d = 0,203$ nm. En abscisses on porte les valeurs de $\sqrt{V}$ et en ordonnées l'intensité de réflexion relative. Sur cette figure des fléchettes indiquent les positions des maximums calculées à l'aide de la formule (18.2). Selon la théorie, les maximums devaient être équidistants avec une distance de séparation égale à $3,06V^{1/2}$. L'expérience confirma ces prévisions mais seulement pour les grandes valeurs de m ($m = 6, 7, 8$). Pour m petit, on constatait des écarts systématiques par rapport à la formule (18.2).

3. La cause de ces écarts fut établie par Bethe (né en 1906) qui démontra que dans le cristal l'indice de réfraction des ondes élec-

*) En fait, la réflexion et la diffusion des ondes longues (longueur d'onde grande devant les distances intermoléculaires et interatomiques) peuvent être également considérées comme résultant d'interférences. Dans ces cas, le rôle de plans réticulaires est assumé par certaines couches parallèles et équidistantes. Mais ces couches sont macroscopiques (cf. t. IV, §§ 66, 99) et non pas monoatomiques.

troniques de de Broglie devait être plus grand que dans le vide. Effectivement, dans le réseau cristallin des métaux, les positions spatiales des ions positifs et des électrons négatifs sont différentes, en conséquence de quoi tout métal est le siège d'un champ électrique dont le potentiel varie périodiquement d'un point à un autre. Dans une approximation grossière, on peut remplacer le potentiel variable par un potentiel constant V_0 qui est la moyenne spatiale du potentiel vrai. Ce potentiel moyen V_0 est appelé *potentiel interne* du métal. Si on pose égal à zéro le potentiel à l'extérieur du métal, V_0 doit être une grandeur positive pour que les électrons restent dans le métal. Dans ce cas, l'énergie potentielle d'un électron à l'intérieur du métal sera négative et l'électron se trouvera enfermé dans un puits de potentiel de profondeur V_0 constante, sur les parois duquel le potentiel change par saut de la valeur zéro à l'extérieur du puits jusqu'à la valeur constante V_0 dans le puits de potentiel. L'existence d'un potentiel interne du métal justifie l'augmentation de l'indice de réfraction lorsqu'on passe du vide dans le métal.

Considérons un électron extérieur tombant sur le métal. Si cet électron a subi l'action d'une différence de potentiel d'accélération V, sa vitesse est $v_1 \sim \sqrt{V}$. Dans le métal cette vitesse atteint la valeur $v_2 \sim \sqrt{V + V_0}$. En pénétrant dans le métal, la trajectoire de l'électron et l'onde de de Broglie qui lui est associée doivent subir une réfraction. Selon (17.13), l'indice de réfraction relatif du métal pour ce processus est

$$\mu_{21} = \frac{v_2}{v_1} = \sqrt{1 + V_0/V}. \tag{18.3}$$

Omettons l'indice 21, i.e. $\mu \equiv \mu_{21}$ (ne pas confondre avec l'indice de réfraction absolu).

Compte tenu de la réfraction des ondes électroniques de de Broglie, la condition de Bragg-Wulff (18.1) doit s'écrire

$$2d\mu \cos \psi = m\lambda \tag{18.4}$$

cf. t. IV, §§ 33, 61) (la perte d'une demi-onde à la réflexion ne se produit pas ici). ψ désigne ici l'angle de réfraction et φ l'angle de glissement (et non pas l'angle d'incidence). On écrira donc la loi de la réfraction sous la forme

$$\frac{\cos \varphi}{\sin \psi} = \mu.$$

On en tire $\cos \psi = \frac{1}{\mu}\sqrt{\mu^2 - \cos^2 \varphi}$, et la condition de Bragg-Wulff devient

$$2d\sqrt{\mu^2 - \cos^2 \varphi} = m\lambda. \tag{18.5}$$

Les calculs confirment ces explications. Dans les expériences pour un angle φ donné, on mesure les valeurs du potentiel d'accélération

V auxquelles correspondent les maximums de Bragg de différents ordres ($m = 3, 4, 5$). A l'aide de la formule (17.16), on évalue les longueurs des ondes de de Broglie correspondant à ces maximums, puis à l'aide de (18.5) on calcule les indices de réfraction μ. Enfin, par la formule (18.3), on calcule le potentiel interne V_0 du métal. On constate qu'aux erreurs expérimentales près, V_0 ne dépend que de la nature du métal et ne dépend pas de l'ordre de réflexion m (i. e. de V) et de l'angle φ. C'est ce que l'on doit avoir si l'interprétation donnée plus haut est juste. Pour le nickel par exemple, $V_0 \approx 15$ V. Les potentiels internes des autres métaux sont du même ordre de grandeur. Notons que pour expliquer certains phénomènes, on ne peut plus remplacer le potentiel vrai du métal par un potentiel moyen, car il faut tenir compte de la variation périodique spatiale du potentiel vrai. C'est pour en tenir compte que Bethe élabora, en 1928, une *théorie dynamique des interférences des ondes de de Broglie*.

4. La méthode qui a permis d'obtenir ces résultats a été dénommée *méthode de Bragg*. L'étude de la diffraction des rayons X peut être réalisée aussi par d'autres méthodes — la *méthode de Laue* (1879-1960) et la *méthode de Debye-Scherrer* (1890-1970)-*Hell* (cf. t. IV, § 61). On peut également utiliser ces méthodes pour l'étude de la diffraction des ondes de de Broglie. En 1927, Davisson et Germer appliquèrent la méthode de Laue. Dans cette méthode, on fait passer un pinceau de rayons X à travers un monocristal. Les rayons diffractés tombent sur une plaque photographique et y sont enregistrés sous la forme de taches distribuées de façon symétrique. Il a été établi (voir t. IV, § 61) que dans ce cas le spectre des rayons X doit être *nécessairement continu*. Dans l'étude de la diffraction des électrons par la méthode de Laue, on devrait utiliser des faisceaux d'électrons ayant des vitesses continuellement variables. Or cela est pratiquement irréalisable, vu que les vitesses des électrons sortant du canon électronique sont les mêmes, ou plus exactement ont une distribution continue, mais dans les limites très étroites. Les conditions de Laue (corrigées pour tenir compte de l'indice de réfraction des ondes électroniques de de Broglie) appliquées à un faisceau d'électrons ayant une vitesse donnée, ne sont pas en général vérifiées et il ne se produit pas d'interférences. Pour obtenir des maximums de diffraction, on est obligé de modifier progressivement la vitesse des électrons (en faisant varier le potentiel d'accélération appliqué au canon à électrons) et de modifier simultanément la position du collecteur.

5. Davisson et Germer utilisèrent un monocristal de nickel (de ymétrie cubique) taillé parallèlement au plan d'indice de Miller (III) (voir t. II, § 135). Ils envoyaient dans le vide, normalement à ce plan, un faisceau d'électrons produit par un canon à électrons (fig. 32). Pour détecter les électrons diffusés, ils utilisaient un double cylindre de Faraday, les cylindres extérieur et intérieur étant isolés

l'un de l'autre par du quartz. Le cylindre intérieur était connecté à un galvanomètre sensible. Pour éliminer les électrons diffusés de façon non élastique et ayant perdu une partie notable de leur énergie, un potentiel de freinage (négatif) était appliqué au cylindre extérieur ; ce potentiel négatif était dix fois plus petit que le potentiel d'accélération. Ainsi, seuls les électrons ayant subi une réflexion élastique, i.e. ceux dont la vitesse a été réduite de moins de 10 %, parvenaient au collecteur. On déterminait le nombre des électrons diffusés en fonction de : 1) l'angle de diffusion, 2) la vitesse des électrons dans le faisceau, 3) l'azimut, i.e. l'angle de rotation du cristal autour d'un axe perpendiculaire au plan de diffusion.

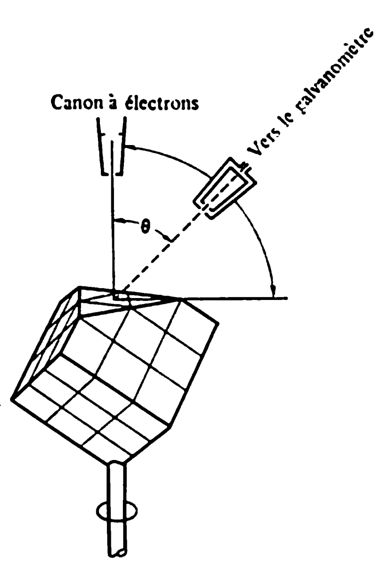


Fig. 32

La figure 33 représente le diagramme polaire caractérisant le nombre d'électrons diffusés en fonction de l'angle de diffusion pour la position du cristal, indiquée sur la figure 32. Dans cette position, la face taillée du cristal est recouverte de rangées d'atomes régulières, perpendiculaires au plan d'incidence (plan de la figure) et se trouvant les unes des autres à une distance $d = 0{,}215$ nm. Le collecteur pou-

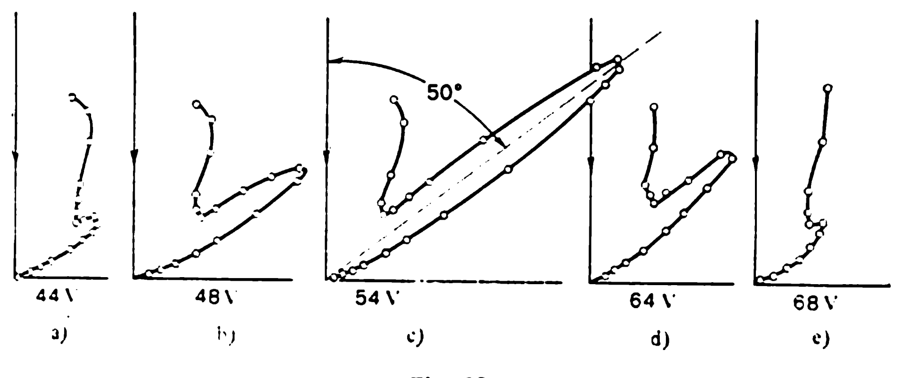


Fig. 33

vait être tourné en restant toujours dans le plan d'incidence. Le potentiel d'accélération variait lui aussi. Pour un angle $\theta = 50°$

et pour différentes tensions d'accélération, on observait un maximum qui ne se développait pleinement que lorsque la tension d'accélération atteignait 54 V (λ = 0,167 nm). On peut interpréter ce maximum comme un maximum de diffraction de premier ordre produit par un réseau de diffraction plan ayant une période d = 0,215 nm. En effet, en calculant l'angle θ par la formule $d \sin \theta = \lambda$, on trouve $\theta = 51°$, en bon accord avec l'expérience.

Dans les expériences, on déterminait aussi les nombres d'électrons diffusés pour différents azimuts du cristal, la position du col-

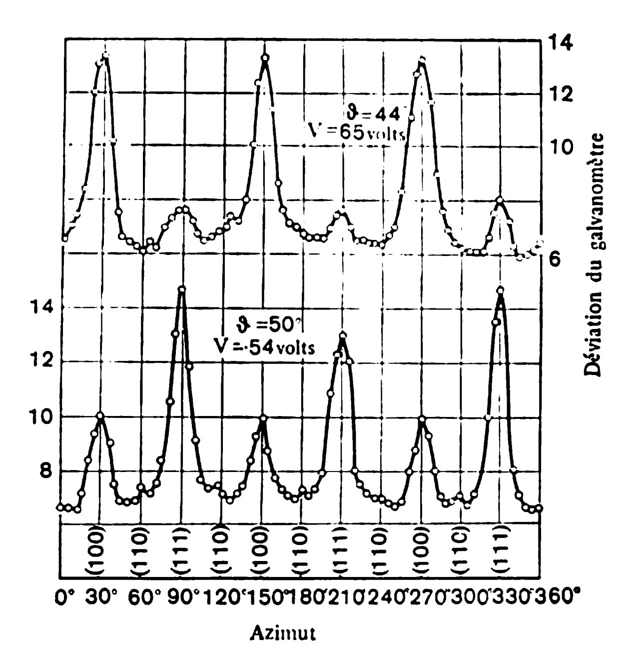


Fig. 34

lecteur et l'angle θ étant maintenus constants. Autrement dit, on faisait tourner le cristal autour d'un axe perpendiculaire au plan taillé. Les résultats de ces mesures sont représentés sur la figure 34 pour θ = 44° et θ = 50° et des tensions d'accélération correspondant à ces angles. La majorité des maximums que l'on discerne sur ces figures correspondent à des interférences spatiales, i.e. à des interférences mutuelles des ondes diffusées par les atomes du réseau spatial du cristal. Lorsqu'on fait tourner le cristal de 360°, ces courbes se reproduisent trois fois. La raison en est que la perpendiculaire au plan réfléchissant (III) est un axe de rotation de troisième ordre du cristal. Pour arriver à un accord quantitatif complet avec la théorie,

il fallait tenir compte de l'indice de réfraction du cristal. Un petit nombre de maximums correspondent au renforcement interférentiel des ondes diffusées par des atomes adsorbés à la surface du cristal. C'est pour cela que dans la partie inférieure de la figure 34 le maximum central est moins haut que les deux autres grands maximums.

6. Les études de la diffraction des électrons par la méthode de Debye-Scherrer-Hell ont été commencées en 1928 par G. P. Thomson

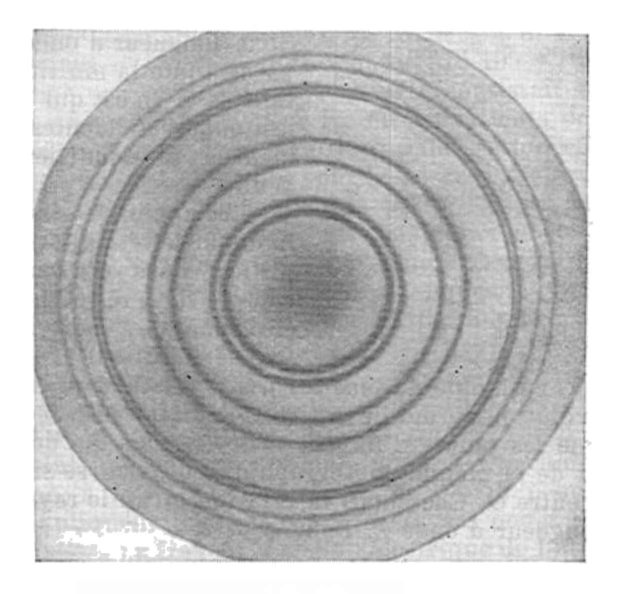


Fig. 35

(1892-1975), P. S. Tartakovski (1895-1939) et quelques autres physiciens. Thomson faisait passer un pinceau monochromatique d'électrons rapides (potentiel d'accélération de 17,5 à 56,5 kV; les longueurs d'onde correspondantes, calculées par la formule relativiste (17.16a), étaient comprises entre 0,092 et 0,052 nm) à travers une feuille polycristalline de $\sim 10^{-5}$ cm d'épaisseur. (L'utilisation d'électrons rapides tient à ce que les électrons plus lents sont fortement absorbés par les feuilles métalliques.) On obtenait sur une plaque photographique, disposée derrière la feuille métallique, une tache centrale entourée d'anneaux de diffraction. La figure 35 représente le diagramme de diffraction obtenu par ce procédé avec une feuille d'or. Pour démontrer simplement que la figure de diffraction avait été produite par les électrons diffusés eux-mêmes et non par les rayons X secondaires qu'ils auraient excités, Thomson appliquait un champ magnétique sous l'action duquel la figure de diffraction était

déplacée et déformée. Si la figure avait été produite par la diffraction des rayons X, le champ magnétique ne l'aurait pas déformée.

L'origine des anneaux de diffraction est ici la même que dans le cas de la diffraction des rayons X (cf. t. IV, § 61). La feuille polycristalline que traverse le faisceau d'électrons se compose d'une multitude de très petits cristallites ($\sim 10^{-6}$ cm) orientés au hasard. Il a été indiqué ci-dessus que pour appliquer la méthode de Debye-Scherrer-Hell, le rayon électronique devait être monochromatique. Si la longueur d'onde λ est fixée, parmi toute la multitude de cristallites il en est qui réfléchissent les ondes incidentes, conformément à la condition de Bragg-Wulff $2d \sin \varphi = m\lambda$ (fig. 36), où φ est l'angle de glissement et d la distance réticulaire. (Pour simplifier, nous n'avons pas tenu compte de la réfraction des rayons électroniques, mais le lecteur le fera sans difficulté aucune.) Statistiquement, l'ensemble de tous les petits cristaux présente une symétrie de révolution autour de la direction du faisceau incident ABO. Il s'ensuit que les points d'impact C des rayons qui déterminent le noircissement de la plaque photographique doivent se situer sur des anneaux de centre O. Cherchons la relation entre le rayon r de l'anneau et la longueur d'onde λ. Pour φ petit

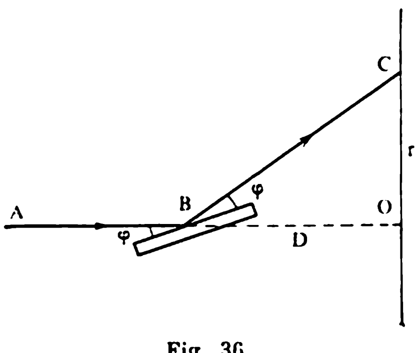


Fig. 36

$$\sin \varphi \approx \varphi = \frac{m\lambda}{2d} = \frac{r}{2D};$$

D étant la distance de l'échantillon à la pellicule. On doit donc avoir

$$\frac{r}{\lambda} = \text{const.} \tag{18.6}$$

Cette relation reste valable lorsqu'on tient compte de la réfraction. En utilisant la formule (17.16) (pour les électrons lents) ou la formule (17.16a) (pour les électrons rapides), (18.6) se transforme en une relation entre r et V. Ces relations ont été confirmées par l'expérience.

Une autre preuve quantitative de la diffraction des électrons par les cristaux est fournie par la comparaison de l'électronogramme et du rœntgénogramme d'un même cristal. Chacun de ces diagrammes de diffraction permet de calculer le paramètre du réseau du cristal; on constate que ces méthodes très différentes conduisent au même résultat.

P. S. Tartakovski entreprit des recherches analogues de la diffraction des électrons en utilisant la méthode de Debye-Scherrer-Hell. Tartakovski utilisa des électrons lents (potentiel d'accélération allant jusqu'à 1700 V, $\lambda > 0,297$ nm) et des feuilles d'aluminium. Les expériences étaient réalisées avec un angle de diffusion constant. Lorsqu'on faisait varier la vitesse des électrons, on observait une série de maximums pour lesquels la condition de Bragg-Wulff était vérifiée.

7. On peut observer la diffraction des électrons par des atomes et des molécules isolés. Sans entrer dans les détails de cette question, notons seulement qu'en 1920 Ramsauer (1879-1955) fit une découverte qui porte son nom. Cette découverte consistait en ce que la perméabilité de certains gaz pour les électrons lents présentait une augmentation anormale. Cela signifie que la section efficace totale d'un atome (cf. § 9 et t. II, § 87) pour le faisceau d'électrons passant à travers le gaz devient anomalement petite lorsque la vitesse des électrons diminue. Ce phénomène est particulièrement net dans les gaz rares — argon, crypton et xénon, et ne s'observe pas dans l'hélium et le néon. Selon les conceptions classiques (cf. § 9), la section efficace des atomes devrait diminuer de façon monotone à mesure que la vitesse des électrons augmente. Mais l'expérience montre que pour l'argon, le crypton et le xénon, la section efficace présente un minimum important pour une tension d'accélération $\sim 0,6$ V. De part et d'autre de ce minimum, la section efficace croît brusquement et atteint une valeur maximale pour une tension égale à 6 V pour le xénon, à 11 V pour le krypton et à 13 V pour l'argon. Ces résultats s'expliquent par les interférences des ondes de de Broglie lors de leur passage à travers le champ électrique de l'atome. On peut dire *grosso modo* que l'influence de ce champ peut être ramenée à l'action d'un milieu non homogène, dont l'indice de réfraction varie de façon continue d'un point à un autre dans l'atome.

8. Jusqu'ici il n'a été question que de la diffraction des électrons. Les atomes et les molécules sont également susceptibles de diffraction. Stern (1888-1969) montra que dans le cas d'atomes lourds pour lesquels la longueur de l'onde de de Broglie $\lambda = h/(mv)$ est très petite, la figure de diffraction n'apparaît pas, ou si elle apparaît elle est floue. Avec les atomes et les molécules légers (H_2, He), on observe des figures de diffraction fort nettes satisfaisant avec une grande précision à la relation $\lambda = h/(mv)$. L'application des formules (17.18) et (17.19) montre que pour la molécule d'hydrogène et l'atome d'hélium la longueur d'onde moyenne est, à la température ordinaire, de l'ordre de 0,1 nm, i.e. comparable au paramètre du réseau cristallin. Comme les vitesses de ces particules présentent une forte dispersion, l'obtention de bonnes figures de diffraction n'est possible qu'en rendant monochromatiques les faisceaux de ces particules. La figure 37 représente le schéma d'un sélecteur des vites-

ses des molécules. Ce dispositif se compose de deux disques à fentes, emmanchés sur un axe commun; les fentes des deux disques sont identiques, mais sont tournées d'un certain angle les unes par rapport aux autres. Les disques peuvent tourner autour de l'axe. Un faisceau d'atomes ou de molécules est envoyé sur les disques parallèlement à l'axe de rotation. En principe, le sélecteur fonctionne à l'instar de la roue dentée de l'appareil bien connu de Fizeau (1819-1896) servant à mesurer la vitesse de la lumière. En ce qui concerne les figures de diffraction produites par les pinceaux atomiques et moléculaires, on notera que ces particules ne pénètrent pas dans le cristal. Les cristaux se comportent vis-à-vis de ces pinceaux comme des réseaux réflecteurs bidimensionnels plans. La figure 38 représente la figure de diffraction des atomes d'hélium par un cristal de fluorure de lithium, i.e. la distribution angulaire des atomes diffractés.

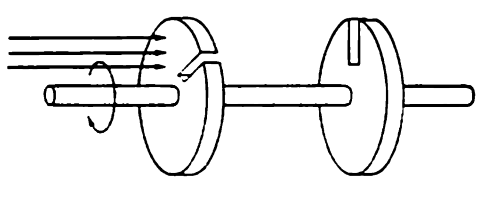


Fig. 37

9. Examinons brièvement la *diffraction des neutrons*. Le neutron est une particule élémentaire neutre dont la masse est presque la même que celle du proton. Dépourvu de charge électrique, le neutron n'interagit qu'avec les noyaux des atomes par l'intermédiaire des forces nucléaires. Le neutron n'exerce aucune action sur les électrons des enveloppes atomiques et n'impressionne donc pas les pellicules photographiques. Pour pouvoir néanmoins utiliser celles-ci, on recouvre la couche photosensible d'une feuille d'indium; les réactions nucléaires se produisent aux points d'impact des neutrons, avec libération d'électrons qui impressionnent l'émulsion. Pour obtenir des figures de diffraction (neutronogrammes) par la méthode de Laue, on envoie un faisceau « blanc » de neutrons thermiques sortant d'un réacteur nucléaire, sur un gros monocristal, dans lequel se produit la diffraction des neutrons. La figure 39 représente un neutronogramme d'un monocristal de NaCl. En plus de la tache centrale, on y voit un ensemble de taches symétriques correspondant

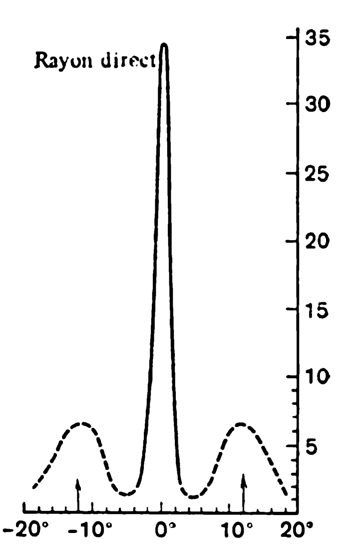


Fig. 38

à l'axe de rotation de 4e ordre du cristal de NaCl. On notera que les taches sont isolées. Comme indiqué dans le t. IV, § 61, cela résulte de ce que la plaque photographique n'a pas été placée en position de diffraction de Fraunhofer, mais dans une position proche du cristal

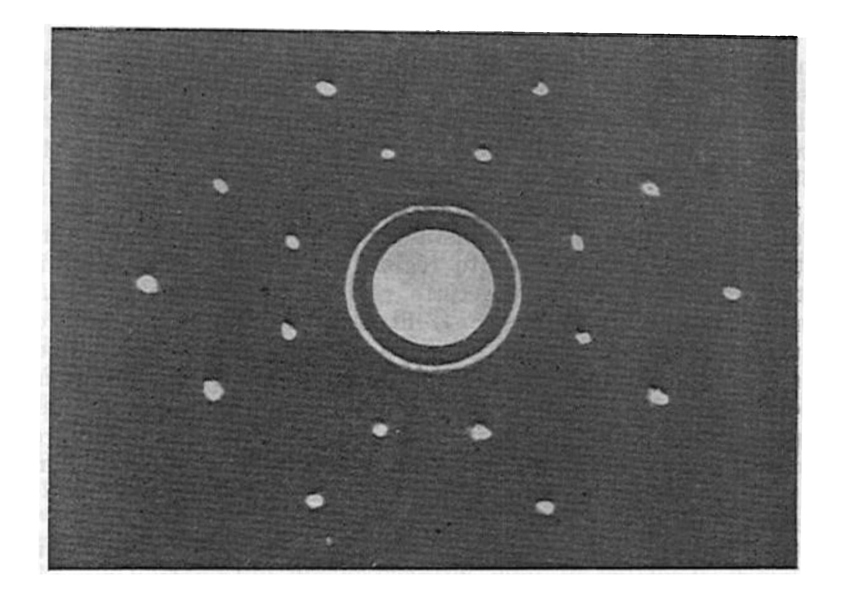


Fig. 39

où *s'applique l'optique géométrique*; cependant la distance au cristal est encore assez grande pour que les différents pinceaux formés dans le cristal à partir de l'onde incidente se soient séparés.

Pour résumer, on dira que les propriétés ondulatoires des particules ont été confirmées par l'expérience et trouvent de nombreuses applications dans les sciences et les techniques (électronographie, neutronographie, etc.).

PROBLÈME

Un faisceau d'atomes ou de molécules sortant d'un four, où leurs vitesses sont régies par la loi de Maxwell, est envoyé sur un sélecteur de vitesses (fig. 37), qui sélectionne un pinceau de particules dont les vitesses absolues sont comprises entre v et $v + dv$, l'intervalle dv étant constant. Pour quelle valeur de v (sélectionnée par une vitesse de rotation donnée des disques) le nombre de particules passant au travers du sélecteur sera-t-il maximal?

R é p o n s e. $v^2 = 3/2\,\frac{kT}{m}$, i.e. v doit être la vitesse quadratique moyenne.

§ 19. Interprétation statistique des ondes de de Broglie et de la fonction d'onde

1. Voyons quelle est la signification physique des ondes de de Broglie et quels sont leurs liens avec les particules de matière.

Schrödinger fut pendant quelque temps partisan de l'idée suivante avant de l'abandonner définitivement. La dualité onde-corpuscule n'existerait pas dans la Nature, seules les ondes seraient réelles, tandis que les particules résulteraient d'une *superposition d'ondes*. En effet, selon le théorème mathématique de Fourier (1768-1830), en composant des ondes de différentes fréquences et directions, on obtient un *paquet d'ondes*, i.e. une construction ondulatoire telle qu'en se superposant à un instant donné les ondes se renforcent mutuellement dans une petite région de l'espace et se détruisent mutuellement en dehors de cette région. *Schrödinger identifia ce paquet d'ondes à la particule.* L'intensité des ondes de de Broglie est alors considérée comme une grandeur proportionnelle à la densité du milieu à partir duquel se forme la particule. Il pouvait sembler que ce point de vue était justifié par ce que le centre du paquet d'ondes doit se propager dans le vide, à l'instar du centre d'un groupe d'ondes, avec la vitesse de groupe. Or, d'après la formule (17.8), la vitesse de groupe des ondes de de Broglie est égale à la vitesse de la particule.

Or un paquet d'ondes ne peut se comporter indéfiniment comme une particule. La raison en est que même dans le vide, les ondes de de Broglie présentent une *dispersion*. Effectivement, la relation entre l'impulsion p de la particule et son énergie $\mathscr{E}$ s'exprime par

$$\left(\frac{\mathscr{E}}{c}\right)^2 - p^2 = (m_0c)^2. \qquad (19.1)$$

En y substituant les valeurs de $\mathscr{E}$ et de p tirées des formules (17.2) et (17.3), on obtient la loi de dispersion des ondes de de Broglie

$$\left(\frac{\hbar\omega}{c}\right)^2 - (\hbar k)^2 = (m_0c)^2. \qquad (19.2)$$

Cette expression montre que la vitesse de phase de ces ondes $v_{\text{ph}} = \omega/k$ dépend de la fréquence ω et c'est ce qui caractérise leur dispersion. Supposons qu'à un instant $t = 0$, les ondes de de Broglie se renforcent mutuellement dans une petite région de l'espace, tandis que dans le reste de l'espace le champ d'onde s'annule. Comment évoluera le paquet d'ondes dans le temps? Les ondes monochromatiques de fréquences différentes qui le composent commenceront à se séparer avec des vitesses de phase différentes, et le paquet d'ondes se déformera, puis il *se dispersera* finalement dans l'espace. Donc, si une particule était un paquet d'ondes, elle serait *instable* et devrait

se décomposer rapidement. Or cela ne correspond nullement à la réalité. Il s'ensuit qu'*une particule ne peut être assimilée à un paquet d'ondes composé d'ondes de de Broglie.*

2. Le point de vue contraire est tout aussi inacceptable: les *particules* sont primaires et les ondes en sont des *produits*, i.e. se forment dans le milieu composé de particules, comme le son se propageant dans l'air. Ce milieu devrait être suffisamment dense, car on ne peut parler d'ondes dans un milieu de particules que si la distance moyenne entre les particules est très petite devant la longueur d'onde. On a vu plus haut que dans les cas les plus typiques, cette condition n'est pas remplie pour les ondes de de Broglie. Mais même si cette difficulté pouvait être surmontée, ce point de vue devrait pourtant être rejeté, car il présuppose que les propriétés d'onde sont propres aux systèmes formés de plusieurs particules et non pas aux particules isolées. Or les propriétés ondulatoires des particules, leur aptitude aux interférences ne s'évanouissent pas lorsque l'intensité des faisceaux incidents devient petite. Cela a été démontré à l'aide d'expériences directes avec des électrons par Biberman, Souchkine et Fabricant (né en 1907), et par Janossi (1912-1978) pour les photons.

Biberman, Souchkine et Fabricant utilisaient des faisceaux comportant un nombre d'électrons tellement petit que l'intervalle de temps moyen entre deux passages successifs de l'électron à travers le système de diffraction était 30 000 fois environ plus long que le temps que mettait un même électron pour le parcourir. Dans ces conditions, toute possibilité d'interaction mutuelle doit être exclue. Néanmoins, si le temps d'exposition était suffisamment long, il se formerait des figures de diffraction qui ne se distingueraient en rien de celles obtenues avec des faisceaux électroniques d'une intensité 10^7 fois plus grande environ et un temps d'exposition court. Ce qui compte, c'est que le nombre total d'électrons qui tombaient sur la pellicule était le même dans les deux cas. Ces expériences montrent que *les particules isolées possèdent des propriétés ondulatoires*; une preuve indirecte en est que les électrons des enveloppes atomiques, par exemple l'unique électron de l'atome d'hydrogène, manifestent des propriétés d'onde, quoiqu'il ne peut alors être question de milieu formé d'électrons.

3. Ces deux tentatives de trouver une interprétation physique aux ondes de de Broglie étaient purement hypothétiques. Confrontées aux données expérimentales, elles s'effondrèrent. Pour ne plus nous engager dans la voie des hypothèses, nous utiliserons une méthode plus sure, *celle des principes.* Soit un faisceau de particules (des électrons, par exemple) tombant sur un dispositif de diffraction (un cristal par exemple). Comme indiqué ci-dessus, les ondes de de Broglie associées à une seule particule sont susceptibles de diffraction. On peut donc supposer que le faisceau incident se compose

d'un seul électron. Lorsque l'onde associée à l'électron traverse le cristal, elle se divise en plusieurs faisceaux diffractés. On ne peut évidemment admettre que dans chacun de ces faisceaux se trouve une portion d'électron. L'électron se manifeste comme un tout et on ne décèle jamais une portion d'électron, conformément à l'atomisme qui est propre au microcosme. Disposons sur le trajet de l'un des faisceaux diffractés un capteur d'électrons. Si ce dispositif se déclenche, il enregistre toujours un *électron entier* et jamais une portion d'électron. Mais on ne peut en conclure qu'avant d'être décelé l'électron se trouvait dans un seul des faisceaux diffractés et que les autres faisceaux ne jouaient aucun rôle dans le phénomène, qu'ils étaient inexistants. Si on adoptait ce point de vue, cela signifierait que l'électron traversait l'appareil comme le ferait un point matériel de la mécanique classique. *Un tel comportement est absolument incompatible avec les phénomènes d'interférences et de diffraction des électrons.* Si on répète l'expérience avec un autre électron, celui-ci sera également décelé dans un des faisceaux diffractés qui, en règle générale, n'est pas le même que pour le premier électron. Faisant face à ces difficultés, Born avança une *interprétation statistique* des ondes de de Broglie, qui permettait de concilier l'atomisme des particules avec les propriétés d'onde.

4. Selon l'interprétation statistique, on doit considérer les ondes de de Broglie comme des *ondes de probabilité* ou plus précisément: *l'intensité de l'onde de de Broglie en un certain endroit de l'espace est proportionnelle à la probabilité de présence de la particule en cet endroit.* Mais pour établir les propriétés statistiques ou probabilistes des particules, l'expérience doit porter non pas sur une seule, mais sur un *grand nombre de particules* ou encore on doit répéter un *grand nombre de fois* l'expérience portant sur une seule particule dans les mêmes conditions. On ne peut parler de statistique et de probabilité qu'ayant en vue un ensemble déterminé d'éléments que concernent ces notions. Ce peut être un ensemble de plusieurs éléments observés simultanément, ou bien un seul élément soumis à maintes observations à des instants consécutifs. En mécanique quantique, ces ensembles sont appelés *ensembles quantiques.* Selon l'un des principes fondamentaux de la mécanique quantique, un *ensemble quantique ne peut être réalisé qu'en fixant certains paramètres macroscopiques.* Cela ne signifie nullement que les propriétés ondulatoires sont propres aux ensembles de particules et non pas aux particules elles-mêmes. Les ensembles ne sont nécessaires que pour *mettre en évidence* les propriétés d'onde.

L'interprétation statistique des phénomènes s'est développée en physique statistique. Mais à la base de la statistique classique on trouve des lois dynamiques qui admettent en principe la possibilité de prédire exactement le comportement dans le temps de chacune des microparticules du système. Il suffirait pour cela de connaître les

conditions initiales pour toutes les microparticules ainsi que les forces d'interaction s'exerçant entre celles-ci. C'est ainsi que le comportement d'un gaz au repos est macroscopiquement caractérisé par sa température, sa pression et son volume. Mais on pourrait, en principe, décrire le comportement du gaz en indiquant à tout instant les positions et les vitesses de toutes les molécules. Par rapport à la description macroscopique, les positions et les vitesses des molécules du gaz jouent le rôle de *paramètres cachés*, i.e. de grandeurs ne se manifestant pas à l'échelle macroscopique.

On pourrait penser qu'à la base de la description statistique des phénomènes, donnée par la mécanique quantique, se trouvent des lois dynamiques encore inconnues, mais qui régissent des paramètres cachés également inconnus, susceptibles de caractériser l'état des systèmes de façon plus approfondie et plus fine que nous ne savons le faire actuellement. C'est un point de vue qui existe réellement, mais jusqu'à présent il n'a conduit à aucun résultat positif. Aussi laisserons-nous de côté la question des paramètres cachés et ne donnerons-nous que la description statistique que fournit de nos jours la mécanique quantique. Il importe de noter que cette description est suffisante pour prévoir et interpréter les résultats expérimentaux, tout au moins au niveau de nos connaissances.

5. Voyons l'explication que fournit la statistique des phénomènes de diffraction des particules, par exemple des électrons. Avant de pénétrer dans l'appareil de diffraction, l'électron est accéléré par une certaine différence de potentiel, à laquelle correspond une longueur donnée de l'onde de de Broglie. Le potentiel d'accélération est le paramètre macroscopique qui détermine un ensemble quantique de particules. Supposons que l'enregistrement des électrons est réalisé à l'aide d'une plaque photographique. Il est impossible de prédire avec certitude le point d'impact d'un électron donné ; on ne peut qu'indiquer la probabilité de cet événement. La probabilité d'impact d'un électron en un point donné de la plaque est proportionnelle à l'intensité de l'onde de de Broglie en ce point. L'impact d'un électron isolé fait apparaître (après révélation) une petite tache sur la plaque. Si le nombre d'électrons est petit, la plaque aura l'aspect d'une cible qui aurait été perforée par un petit nombre de balles. La répartition des taches sur la plaque ne présente aucune régularité. Une régularité dans la répartition des taches n'apparaîtra que lorsqu'un très grand nombre d'électrons seront tombés sur la plaque. Dans ce cas, ils tomberont de préférence là où doivent apparaître les maximums de diffraction des ondes de de Broglie. L'ensemble des taches correspondantes forment la figure de diffraction que fournit l'expérience. Tant que la répulsion électrostatique mutuelle des électrons n'intervient pas, les figures de diffraction que l'on obtient sont les mêmes, qu'elles résultent du passage (un par un) d'un grand nombre d'électrons à travers l'appareil de diffraction

ou du passage d'un intense faisceau d'électrons ayant tous la même vitesse, à condition que le nombre total d'électrons soit le même dans les deux cas.

6. Le caractère symbolique des ondes de de Broglie se manifeste notamment en ce que la fonction (17.1), qui représente une onde plane de de Broglie $\Psi(\boldsymbol{r}, t)$, est *essentiellement complexe.* On a déjà eu affaire à des expressions complexes de cette forme en physique classique, mais on n'attribuait une signification physique qu'à la partie réelle de ces expressions, de sorte qu'on aurait pu tout aussi bien omettre les parties imaginaires. On ne le faisait que pour diminuer le nombre des transformations mathématiques et pour faciliter l'interprétation des résultats. En mécanique quantique, la situation est toute autre, car la fonction $\Psi(\boldsymbol{r}, t)$ est *complexe en principe.* Toutes les grandeurs physiques ayant une signification physique concrète s'expriment par la fonction complexe $\Psi(\boldsymbol{r}, t)$ *toute entière* et non seulement par sa partie réelle. On pourrait fort bien éviter l'utilisation des fonctions complexes, mais dans ce cas on devrait utiliser *deux* fonctions réelles pour représenter l'onde de de Broglie. L'une de ces fonctions représenterait la partie réelle et l'autre représenterait le coefficient de la partie imaginaire de la fonction complexe Ψ. On ne ferait que compliquer inutilement les calculs.

Maintenant, la probabilité de présence de la particule en un point donné de l'espace peut être représentée par le carré du module de la fonction Ψ au même point, i.e. $|\Psi|^2 = \Psi^*\Psi$. Pour une onde de de Broglie plane (17.1), le carré du module est égal à

$$\Psi^*(\boldsymbol{r}, t)\,\Psi(\boldsymbol{r}, t) = \Psi_0^*\Psi_0 = \text{const},$$

ce qui signifie que la probabilité de présence de la particule est la même en tout point de l'espace. Dans le cas d'une particule animée d'un mouvement uniforme pendant un temps indéfiniment long, tout autre résultat serait incompatible avec l'homogénéité de l'espace. Si au lieu d'exprimer l'onde de de Broglie par l'expression complexe (17.1) on avait utilisé une expression réelle, par exemple les fonctions sinus ou cosinus, on n'aurait pas obtenu ce résultat. Cette remarque est une justification de l'utilisation des expressions complexes. On doit encore se demander si les interférences d'ondes essentiellement complexes sont possibles. Il est facile de prouver que les interférences sont effectivement possibles. Soient deux ondes de de Broglie représentées par les expressions

$$\Psi_1 = e^{i(kr-\omega t)}, \quad \Psi_2 = e^{i(kr-\omega t-\delta(r))},$$

qui sont telles qu'il existe entre elles une différence de phase $\delta(\boldsymbol{r})$. Par superposition de ces ondes on obtient l'onde $\Psi = \Psi_1 + \Psi_2$. La probabilité de présence de la particule en un point quelconque

de l'espace est alors proportionnelle à

$$\Psi^*\Psi = (\Psi_1^* + \Psi_2^*)(\Psi_1 + \Psi_2) =$$
$$= (\Psi_1^*\Psi_1 + \Psi_2^*\Psi_2) + (\Psi_2^*\Psi_1 + \Psi_1^*\Psi_2) =$$
$$= 2 + (e^{i\delta} + e^{-i\delta}) = 2(1 + \cos\delta).$$

Cette expression contient le terme interférentiel $2\cos\delta(\boldsymbol{r})$ qui varie dans l'espace de -2 jusqu'à $+2$ en fonction de la différence de phase $\delta(\boldsymbol{r})$. On doit donc pouvoir observer des interférences.

7. Jusqu'ici il n'a été question que d'ondes de de Broglie planes représentées par l'expression (17.1). Les ondes de ce type sont associées aux particules animées d'un mouvement uniforme libre. Il importe de généraliser les résultats déjà obtenus aux cas de *mouvements variés des particules dans des champs de forces arbitraires.* Dans ces cas, pour décrire complètement l'état d'une particule, on utilise en mécanique quantique non pas une onde plane de de Broglie, mais une fonction complexe $\Psi(\boldsymbol{r}, t)$ qui dépend des coordonnées et du temps. Cette fonction est appelée *fonction d'onde.* Dans le cas particulier du mouvement libre de la particule, la fonction d'onde se ramène à l'expression (17.1) de l'onde plane de de Broglie. La fonction d'onde elle-même n'est introduite qu'en qualité de *symbole auxiliaire* et ne fait pas partie des grandeurs directement observables. La connaissance de cette fonction permet cependant de *prédire statistiquement* les grandeurs observables et ayant donc une signification physique réelle. Les liens existant entre la fonction d'onde et les grandeurs observables apparaîtront au fur et à mesure des développements ultérieurs. Ici nous n'examinerons qu'un seul aspect de cette question.

A l'aide de la fonction d'onde on détermine la *probabilité relative* de présence de la particule en différents points de l'espace. Au stade où il n'est question que de rapports des probabilités, la fonction d'onde n'est, par principe, déterminée qu'à un facteur constant arbitraire près. Si on multiplie, en tous les points de l'espace, la fonction d'onde par un même nombre constant (en général par un nombre complexe) différent de zéro, on obtient une nouvelle fonction d'onde qui décrit exactement *le même état.* Il est dénué de sens de dire que la fonction Ψ est égale à zéro en tous les points de l'espace, car une telle « fonction d'onde » ne permettra en aucun cas de tirer des conclusions concernant les probabilités relatives de présence de la particule en différents points de l'espace. Mais on peut réduire notablement l'indétermination sur Ψ *en passant des probabilités relatives aux probabilités absolues.* Arrangeons le facteur indéterminé de la fonction Ψ de telle sorte que la quantité $|\Psi|^2\,dV$ représente la *probabilité absolue* de présence de la particule dans un élément de volume dV de l'espace. Dans ces conditions, $|\Psi|^2 = \Psi^*\Psi$ aura le sens d'une *densité de probabilité* que l'on peut espérer avoir pour

déceler la particule dans l'espace. Mais Ψ sera encore déterminé *à un facteur complexe constant près, dont le module est cependant égal à l'unité*. Cette définition exige que soit vérifiée la *condition de normalisation*

$$\int |\Psi|^2 dV = 1, \tag{19.3}$$

où l'intégrale est étendue à tout l'espace infini. Cette condition implique la certitude de présence de la particule dans la totalité de l'espace.

Si l'intégrale (19.3) est *divergente*, la normalisation est irréalisable. C'est ce qui se produit dans le cas d'une onde de de Broglie plane lorsque la probabilité de présence de la particule est la même en tous les points de l'espace. Mais les cas semblables doivent être considérés comme une idéalisation de la situation réelle où la particule, au lieu de s'éloigner à l'infini, est astreinte à rester dans une région limitée de l'espace. Dans ce dernier cas, la normalisation (19.3) ne présente aucune difficulté. Mais par un passage à la limite on peut obtenir une normalisation rationnelle, même pour une particule qui n'est pas localisée dans une région finie de l'espace (cf: § 30).

8. Nous voyons ainsi qu'une signification physique concrète est attachée non pas à la fonction Ψ mais à son *module* $\Psi^*\Psi$ (ou d'une façon générale à des expressions bilinéaires en Ψ^* et Ψ). Pourquoi s'attache-t-on, en mécanique quantique, à opérer avec des fonctions d'onde Ψ au lieu d'avoir affaire directement aux grandeurs observables $\Psi^*\Psi$? Parce que c'est indispensable pour pouvoir interpréter les propriétés ondulatoires de la matière — les phénomènes d'interférences et de diffraction. La situation est exactement la même que dans toute autre théorie ondulatoire. La théorie ondulatoire admet (en tout cas dans l'approximation linéaire) le *principe de superposition des champs d'onde eux-mêmes* et non pas de leurs intensités afin d'inclure dans la théorie les phénomènes d'interférences et de diffraction des ondes. De même, la mécanique quantique admet, comme l'un de ses postulats de base, le *principe de superposition des fonctions d'onde. Si* $\Psi_1(r, t)$ *et* $\Psi_2(r, t)$ *sont les fonctions d'onde caractérisant deux états quelconques d'une particule, toute combinaison linéaire de ces fonctions à coefficients constants* $c_1\Psi_1 + c_2\Psi_2$ *représentera une fonction d'onde de cette même particule qui caractérisera un de ses états.* Ayant déterminé Ψ par ce procédé, on peut déterminer ensuite la densité de probabilité $\Psi^*\Psi$ de l'état Ψ. Le principe de superposition trouve sa justification dans l'accord de ses conséquences avec les résultats expérimentaux. On n'a pas encore élucidé si le principe de superposition est une loi exacte de la Nature ou s'il n'est valable que dans l'approximation linéaire. Pour le développement ultérieur de la théorie, on admettra que le principe de superposition est exact.

Notons une nouvelle fois, quoiqu'on en parle dans ce qui précède, que la signification physique de la fonction d'onde Ψ est liée non seulement à son module, mais aussi à sa phase qui est définie par la partie imaginaire de cette fonction. S'il s'agissait de la fonction d'onde d'un seul état, on pourrait se satisfaire de la connaissance du module. Mais dès qu'il est question de la superposition d'états, apparaissent leurs interférences; or celles-ci dépendent de la différence relative des phases des fonctions d'onde décrivant les états considérés.

9. Ces considérations permettent de comprendre pourquoi il fallait *démontrer* l'invariance de la phase de l'onde $\omega t - kr$ lorsqu'il s'agissait d'ondes électromagnétiques, par exemple (cf. t. IV, § 107), tandis que dans le cas des ondes de de Broglie *on ne le démontre pas* puisque la fréquence ω est spécialement choisie de manière à assurer cette invariance (cf. § 17, 2). La raison en est que la preuve donnée dans le t. IV (§ 107) présupposait la *possibilité de dénombrer les ondes* défilant devant l'observateur. *Cette possibilité n'existe pas pour les ondes de de Broglie* puisque l'observateur ne peut pas observer l'onde Ψ elle-même, mais seulement la densité de probabilité $\Psi^*\Psi$. C'est pour cela que la preuve mentionnée est inapplicable aux ondes de de Broglie.

10. La fréquence ω des ondes de de Broglie et, d'une façon générale, la fréquence de la fonction d'onde sont des *grandeurs non observables par principe*. On met ce fait à profit pour élaborer la *forme non relativiste* de la mécanique quantique. En mécanique classique, l'approximation non relativiste permet d'embrasser une large gamme de phénomènes. Il en est de même en mécanique quantique. D'ailleurs, le passage à la forme relativiste est rendu compliqué pour les raisons suivantes. Dans les champs forts, lorsque l'énergie du champ (par exemple du quantum gamma) est supérieure à $2m_ec^2$, la *production de paires électron-positon* commence. Dans les cas analogues, on observe la même chose pour les autres particules. C'est pour cela qu'*une mécanique quantique relativiste conséquente ne peut être la théorie d'un seul corps* (d'une seule particule). *La théorie d'un seul corps ne peut être élaborée que dans l'approximation non relativiste.* Dans ce qui suit, nous n'envisagerons que la mécanique quantique non relativiste.

En mécanique quantique non relativiste, on continuera à utiliser les relations

$$\mathcal{E} = \hbar\omega, \qquad (19.4)$$

$$p = \hbar k. \qquad (19.5)$$

Mais on ne tiendra plus compte de l'énergie propre m_0c^2 des particules. Cela signifie qu'à partir d'ici on introduira une nouvelle fréquence qui se distinguera de la fréquence utilisée plus haut d'une quantité constante. Pour cette nouvelle fréquence, on gardera

l'ancienne notation ω. Dans le cas particulier du mouvement libre, $\mathcal{E} = p^2/2m$, et la loi de dispersion devra s'écrire

$$\omega = \frac{\hbar}{2m} k^2, \tag{19.6}$$

au lieu de (19.2). En conséquence, la vitesse de phase des ondes de de Broglie sera donnée par l'expression

$$v_{\text{ph}} = \frac{\omega}{k} = \frac{\hbar k}{2m} = \frac{p}{2m} = \frac{v}{2}. \tag{19.7}$$

Mais ceci ne peut affecter les conclusions physiques de la théorie puisque *la vitesse de phase, ainsi que la fréquence* ω *elle-même des ondes de de Broglie, sont des grandeurs inobservables par principe.* Ce qui compte, c'est que les grandeurs physiquement observables — la densité de probabilité $\Psi^*\Psi$ et la vitesse de groupe (17.8) ne changent pas du fait du nouveau choix de la fréquence (cf. § 17, 3). On verra par la suite que toutes les grandeurs mesurables ne changent pas non plus.

11. Pour conclure, donnons une caractéristique générale de toute théorie physique quantitative, ce qui est particulièrement utile pour l'étude de la mécanique quantique. Ainsi que le notait L. I. Mandelstam (1879-1944), toute théorie comporte en général deux parties complémentaires. La première est l'*appareil mathématique de la théorie*, i.e. les équations reliant entre eux les divers symboles mathématiques de la théorie. L'autre partie de la théorie cherche à préciser les *liens entre ces symboles et la Nature*, i.e. la réalité objective. Sans cette seconde partie, la théorie est illusoire puisqu'il ne restera que des mathématiques pures. Mais si c'est la première partie qui manque, il n'y aura pas de théorie quantitative. *Ce n'est que l'ensemble de ces deux parties qui peut constituer une théorie physique quantitative.*

Toutes les théories classiques débutaient par la seconde partie. La signification des symboles qu'utilisait la théorie était connue d'avance ou était précisée en cours de l'élaboration de la théorie. Ce fut le cas pour le temps, la longueur, la masse, la force, la charge électrique, les intensités des champs électrique et magnétique, etc. Reconnaissons que bien souvent les définitions scientifiques rigoureuses des diverses notions cédaient la place à des conceptions naïves, puisées dans la vie courante, et cela créait bien des difficultés. C'est ce qui se passa pour les concepts de temps et d'espace; la révision critique de ces concepts donna naissance à la théorie de la relativité. Mais l'ordre de construction des théories restait toujours le même: d'abord élaboration des concepts, puis établissement des équations correspondantes. Pour la mécanique quantique, les choses se passèrent autrement. On établit tout d'abord les équations pour plusieurs symboles (la fonction d'onde notamment) dont la signifi-

cation physique restait encore imprécise, et ce n'est qu'après qu'on s'occupa d'établir des liens entre ces symboles et la réalité objective. Quoique ce procédé de construction de la théorie puisse paraître contre nature, il est logiquement justifié si les rapports avec la réalité sont bien établis. C'est ce procédé que nous adopterons. Jusqu'ici nous avons établi un rapport entre la fonction d'onde Ψ et la densité de probabilité correspondante $\Psi^*\Psi$, sans pour cela en avoir fini avec l'interprétation physique de tous les concepts et de toutes les grandeurs qu'on a dû introduire avec la fonction Ψ.

§ 20. Les relations d'incertitude

1. En mécanique classique, l'état d'un point matériel est caractérisé à tout instant par sa *position* et son *impulsion*. Les particules réelles — électrons, protons, atomes, molécules, etc., sont des objets plus compliqués. *On ne peut caractériser l'état instantané d'une microparticule en indiquant de façon précise sa position et son impulsion.* La raison en est que toute microparticule possède des propriétés corpusculaires et ondulatoires. On ne peut pas dire qu'en un point donné de l'espace la longueur d'onde est égale à λ, si on ne sait rien du champ d'ondes régnant dans tous les autres points de l'espace. La longueur d'onde est une *caractéristique de la sinusoïde* et celle-ci est une *courbe périodique infinie.* Si on découpe une petite portion de sinusoïde et qu'on éloigne ses autres parties, la portion restante ne possédera plus la propriété la plus marquante de la sinusoïde — sa périodicité. Pour une portion de sinusoïde petite par rapport à λ, la notion de longueur d'onde n'existe plus. Il est par suite évident que les expressions telles que « la longueur d'onde en un point x de l'espace est égale à λ » ou « la fréquence d'un processus ondulatoire à un instant donné t est égale à ω » sont complètement dénuées de sens puisque la grandeur λ n'est pas une fonction de x et ω n'est pas une fonction de t.

D'autre part, si une formation ondulatoire occupe une région limitée de l'espace, on peut toujours la représenter à l'aide de plusieurs sinusoïdes, car une seule sinusoïde ne saurait suffire. Il faut obtenir un paquet d'ondes, i.e. une superposition d'un grand nombre de sinusoïdes de fréquences différentes qui s'additionneraient dans une région déterminée de l'espace et se détruiraient mutuellement hors de cette région. Si la longueur du paquet d'ondes est égale à Δx (pour simplifier nous considérons le cas unidimensionnel), les nombres d'onde k nécessaires à la formation du paquet d'ondes ne peuvent recouvrir un intervalle Δk aussi petit qu'on le voudrait. La largeur minimale de l'intervalle Δk doit satisfaire de façon approchée à la relation

$$\Delta x \cdot \Delta k \gtrsim 2\pi \qquad (20.1)$$

(cf. t. IV, § 30). C'est là une *relation purement ondulatoire*. On en déduit qu'un signal radio court (Δx petit) comporte toujours, avec des intensités notables, des ondes monochromatiques ayant des valeurs différentes de λ. Ces signaux pourront être captés par des récepteurs accordés sur plusieurs longueurs d'onde. Si on désire recevoir des signaux presque monochromatiques ($\Delta\lambda$ petit), ils seront nécessairement longs (Δx grand).

Rappelons comment on peut parvenir à la relation (20.1) à l'aide de considérations simples. Considérons un grand nombre de sinusoïdes de même amplitude, dont les nombres d'onde k augmentent systématiquement d'une même quantité quand on passe d'une sinusoïde à la suivante. Au point x, les phases des ondes varient depuis la valeur kx jusqu'à la valeur $(k + \Delta k)\, x$, i.e. la variation est égale à $x\, \Delta k$. Si $x\, \Delta k = 2\pi$, à ce point toutes les sinusoïdes se détruisent mutuellement. Déterminons le point le plus proche $x + \Delta x$ où les sinusoïdes se détruisent mutuellement. A ce point, la différence de phase entre les sinusoïdes extrêmes est égale à

$$(k + \Delta k)\,(x + \Delta x) - k\,(x + \Delta x) = x\,\Delta k + \Delta x \cdot \Delta k = 2\pi + \Delta x \cdot \Delta k.$$

L'extinction la plus proche aura lieu lorsque $\Delta x \cdot \Delta k = 2\pi$. Il s'ensuit que la totalité de la perturbation ondulatoire peut être divisée en tronçons de longueur Δx aux extrémités desquels le champ d'ondes sera nul. On a obtenu ce résultat parce que toutes les sinusoïdes ont même amplitude. Dans le cas où on considérerait des sinusoïdes d'amplitudes différentes, on pourrait amplifier la perturbation dans les limites d'un seul tronçon Δx et l'annuler partout en dehors de ce tronçon. Ce résultat découle du théorème mathématique de Fourier avec une condition nécessaire de la forme (20.1). C'est ce cas que nous aurons en vue dans ce qui suit.

2. Considérons maintenant un paquet d'ondes de de Broglie dont l'étendue et les limites correspondantes des nombres d'onde vérifient la condition (20.1). Conformément à l'interprétation statistique, la probabilité de présence de la particule ne sera différente de zéro que dans les limites du paquet d'ondes. Quelle sera l'impulsion de la particule? A chaque onde de de Broglie de vecteur d'onde $\mathbf{k}$ correspond une impulsion $\mathbf{p} = \hbar \mathbf{k}$. Le paquet d'ondes ne possède pas d'impulsion déterminée; on a un ensemble d'impulsions occupant l'intervalle compris entre $\mathbf{p} = \hbar \mathbf{k}$ et $\mathbf{p} + \Delta \mathbf{p} = \hbar\,(\mathbf{k} + \Delta \mathbf{k})$ et on ne peut dire quelle valeur de l'impulsion fournira la mesure. Tout ce que l'on peut faire, c'est indiquer les probabilités des résultats de mesure. Par mesure on trouvera avec différentes probabilités des valeurs de l'impulsion comprises entre $\mathbf{p} = \hbar \mathbf{k}$ et $\mathbf{p} + \Delta \mathbf{p} = \hbar\,(\mathbf{k} + \Delta \mathbf{k})$. En exprimant $\mathbf{k}$ en fonction de $\mathbf{p}$, la relation (20.1) s'écrira

$$\Delta x \cdot \Delta p \gtrsim 2\pi\hbar = h. \qquad (20.2)$$

Cette relation est appelée *relation* ou *principe d'incertitude de Heisenberg pour la coordonnée et l'impulsion d'une particule.* Elle détermine les *limites des incertitudes* sur les valeurs de Δx et Δp qui permettraient de caractériser l'état de la particule selon la théorie classique, i.e. par la coordonnée x et l'impulsion p. Plus la valeur de x est précise, plus la valeur de p sera imprécise et inversement. En aucun cas on ne doit interpréter la relation de Heisenberg en disant qu'à tout instant la particule possède des valeurs déterminées de x et p, mais que, par principe, on ne peut les déterminer avec une précision supérieure à celle que tolère la relation d'incertitude (20.2). Ce point de vue agnostique a été avancé, mais il ne correspond nullement à la nature des micro-objets considérés ici. Le sens exact de (20.2) reflète le fait qu'il n'existe pas dans la Nature d'états des particules qui soient caractérisés par des valeurs exactement déterminées des deux variables x et p.

Le principe d'incertitude a été formulé par Heisenberg en 1927 et contribua grandement à une interprétation correcte des régularités existant dans le microcosme et à l'élaboration de la mécanique quantique.

Dans des cas particuliers, l'incertitude Δp peut être inexistante ($\Delta p = 0$). C'est le cas, par exemple, d'une onde de de Broglie monochromatique et plane. Mais dans ce cas, en vertu de la relation d'incertitude $\Delta x = \infty$, ce qui signifie qu'on ne peut rien dire sur la position où serait localisée la particule, la probabilité de présence de celle-ci sera alors la même en n'importe quel point de l'espace. Dans le cas contraire où $\Delta x = 0$, on aura $\Delta p = \infty$ et la fonction d'onde se contracte en un point. Si la particule est strictement localisée, on pourra la déceler en un point déterminé (à l'origine des coordonnées par exemple), mais on ne pourra porter sur la valeur de son impulsion qu'une estimation probabiliste. On démontre que dans ce cas toutes les valeurs de l'impulsion sont équiprobables.

3. Dans le cas tridimensionnel, en théorie classique on caractérise la particule par trois coordonnées rectangulaires x, y, z, et par trois impulsions conjuguées p_x, p_y, p_z. Dans ce cas, la relation d'incertitude de Heisenberg s'exprime par trois inégalités

$$\begin{aligned} \Delta x \cdot \Delta p_x &\geqslant h, \\ \Delta y \cdot \Delta p_y &\geqslant h, \\ \Delta z \cdot \Delta p_z &\geqslant h. \end{aligned} \tag{20.3}$$

Les relations d'incertitude n'imposent aucune limitation sur les produits tels que $\Delta x \cdot \Delta p_y$, $\Delta y \cdot \Delta p_z$, qui apparaissent par multiplication des incertitudes sur la position par les incertitudes sur les impulsions conjuguées à une autre coordonnée. Les valeurs de x et p_y, x et p_z peuvent avoir simultanément des valeurs exactes.

4. On peut mettre la relation (20.2) sous une forme quantitative

exacte. On la donnera sans en faire la démonstration. Dans la relation (20.2), les valeurs de Δx et Δp ne sont pas exactement déterminées. Mais n'importe quelle fonction d'onde Ψ permet de déterminer la *valeur moyenne* $\bar{x}$ *de la coordonnée*; la composition spectrale de Ψ permet de trouver la *valeur moyenne* $\bar{p}$. Connaissant ces quantités, on trouve les écarts par rapport aux valeurs moyennes $\Delta x = x - \bar{x}$, $\Delta p = p - \bar{p}$, ainsi que les carrés moyens de ces écarts $\overline{\Delta x^2}$ et $\overline{\Delta p^2}$. La forme exacte de la relation d'incertitude est

$$\overline{\Delta x^2}\cdot\overline{\Delta p^2} \geqslant \frac{\hbar^2}{4}. \tag{20.4}$$

Sauf exception, on n'utilisera pas cette dernière relation, car pour trancher les questions de principe il importe de connaître l'*ordre de grandeur de la quantité* $\Delta x\cdot\Delta p$ et non pas sa valeur exacte.

5. On étendra les relations d'incertitude (20.3) au cas de *corps macroscopiques*, quoiqu'on ne dispose pas de données expérimentales cnofirmant ou infirmant ces relations. La seule justification de l'application des relations d'incertitude à ce cas est la certitude que ces relations ont une validité universelle. On posera donc la question suivante. Quel effet exercerait la relation d'incertitude sur le mouvement d'un corps macroscopique, si elle était effectivement valable dans ce cas? Soit une petite bille de masse $m = 1$ g. Déterminons la position du centre de la bille avec une grande précision $\Delta x = 10^{-8}$ cm, i.e. avec une précision égale à la dimension des atomes. L'incertitude sur la valeur de l'impulsion de la bille sera alors égale à $\Delta p \sim h/\Delta x = 6{,}63\cdot 10^{-19}$ g·cm/s, et l'incertitude sur sa vitesse $\Delta v = \Delta p/m \sim 6{,}63\cdot 10^{-19}$ cm/s. Une telle précision ne peut être atteinte dans aucune mesure, et par suite la mise en évidence des écarts par rapport au mouvement classique, déterminés par la relation d'incertitude, se trouvent au-delà des possibilités expérimentales.

La situation est tout autre lorsqu'on étudie le mouvement d'un électron dans un atome. Peut-on parler du mouvement classique de l'électron sur une orbite de Bohr? Considérons, pour fixer les idées, un atome d'hydrogène et la première orbite de Bohr. Pour que ce mouvement ait un sens, il faut que l'incertitude sur la valeur du rayon Δr soit petite par rapport à la valeur du rayon de l'orbite $r = \hbar^2/(me^2)$. Mais dans ce cas, l'incertitude sur la valeur de l'impulsion radiale serait égale à $\Delta p_r \approx h/\Delta r \gg h/r = 2\pi\hbar/r = 2\pi\,(me^2/\hbar)$, autrement dit, supérieure à l'impulsion $p = me^2/\hbar$ de l'électron. On obtient des résultats analogues pour les autres orbites de Bohr, à condition que le nombre quantique ne soit pas trop grand. Dans ces conditions, la conception du mouvement sur des orbites classiques perd toute signification. C'est pour cela qu'en mécanique quantique on a rejeté, dans la description du mouvement des électrons

dans les atomes, la notion de trajectoire, car dans ce cas cette notion ne correspond à rien. La situation est exactement la même pour les autres particules élémentaires se déplaçant dans de très petites régions de l'espace.

6. La relation (20.2) se manifeste chaque fois qu'on s'applique à mesurer la position exacte ou l'impulsion exacte d'une particule. Il apparaît qu'en améliorant la précision de la mesure de la position d'une particule, on diminue la précision de la mesure de son impulsion, et vice versa. Cette proposition sera illustrée par trois exemples.

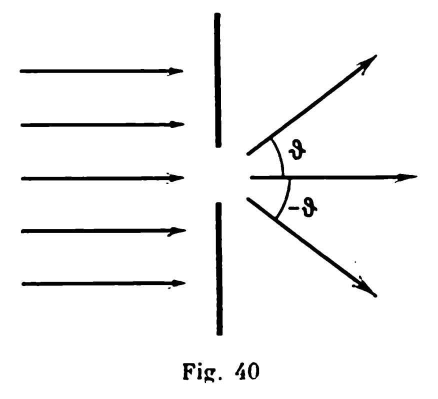


Fig. 40

Premier exemple. Supposons que le mouvement de l'électron soit décrit par une onde monochromatique plane de de Broglie. Dans cet état, l'électron possède une impulsion bien déterminée, mais sa coordonnée est parfaitement indéterminée. Pour déterminer la coordonnée x de l'électron le long du trajet de l'onde, on dispose, perpendiculairement à sa propagation, un écran opaque pourvu d'une fente de largeur d. Faisons coïncider le plan des coordonnées XY avec le plan de l'écran, l'axe X étant perpendiculaire à la fente. Si l'électron passe à travers la fente, dans le plan de la fente la coordonnée x de l'électron sera connue à $\Delta x \sim d$ près. Mais par suite de la diffraction par la fente, la fonction d'onde Ψ de l'électron sera altérée: elle présentera des maximums et des minimums (fig. 40). On peut déceler l'électron partout où $\Psi \neq 0$. Pratiquement, tout le champ d'ondes sera concentré près du maximum central d'ordre zéro. La largeur angulaire de ce maximum est égale à 2ϑ, avec $d \sin \vartheta = \lambda$. En approximation, on peut confondre ce maximum central avec le champ d'ondes en rejetant tout ce qui en reste. Dans cette approximation, après que l'électron aura traversé la fente, l'incertitude Δp_x sur la valeur de son impulsion est de l'ordre de $\Delta p_x = p \sin \vartheta$. En multipliant cette expression par $\Delta x = d$ et en remarquant que $d \sin \vartheta = \lambda$ et $p = h/\lambda$, on trouve $\Delta x \cdot \Delta p_x \gtrsim h$, conformément à la relation de Heisenberg. Plus la fente est étroite, plus précise sera la mesure de la coordonnée de l'électron, mais l'imprécision sur la valeur de l'impulsion ira croissant.

Si on réduit la largeur de la fente jusqu'à $d < \lambda$, on peut rendre l'incertitude sur la position aussi petite que l'on veut. Cependant, pour $d < \lambda$ le champ d'ondes derrière l'écran ne sera plus une onde de de Broglie plane et homogène. On obtient une onde non homogène

qui s'affaiblit rapidement sur une distance de l'ordre de λ ou au-dessous. Dans ce dernier cas, notre estimation de Δp n'est plus valable, mais les études plus poussées montrent que la relation d'incertitude reste valable.

Deuxième exemple (expérience imaginaire de Heisenberg avec un microscope). Soit une particule dans le champ de vision d'un microscope. Pour déterminer sa position, on l'éclaire avec un photon monochromatique de petite longueur d'onde λ (fig. 41). On juge de la position de la particule par le point d'impact du photon sur la plaque photosensible. Si on éclairait avec un faisceau de lumière, l'image de la particule serait une figure de diffraction à anneaux brillants et noirs et une tache claire au centre. Mais on ne peut obtenir une figure de diffraction avec un seul photon. Lorsque le photon est diffusé par l'électron, il peut tomber, avec des probabilités différentes, en n'importe quel point de la figure de diffraction en puissance, à condition que l'intensité lumineuse y soit différente de zéro. Pratiquement, tous ces photons ne tomberont que dans les limites de la tache centrale. On peut négliger l'intensité lumineuse de tous les autres anneaux. Le rayon de la tache centrale est $R \approx \approx \lambda/\beta$. Dans cette approximation, la position du point d'impact du photon dans le plan de l'image peut être déterminée à R près. L'imprécision Δx de la position de l'électron dans le plan objet se déduit de la loi des sinus d'Abbe: $R\beta = \Delta x \sin\alpha$, i.e. $\alpha = \Delta x \sin\alpha$. Plus λ est petit, plus la détermination de la position de la particule sera précise. Mais lorsque le photon est diffusé par l'électron, ce dernier subit un recul et son impulsion s'accroît de façon incontrôlable de $\Delta p_x \sim (h/\lambda) \sin\alpha$. Plus λ est petit, plus cet accroissement incontrôlable de l'impulsion est grand. Ainsi, en effectuant une mesure simultanée de x et p_x, on aboutit à la relation d'incertitude $\Delta x \cdot \Delta p_x \sim h$.

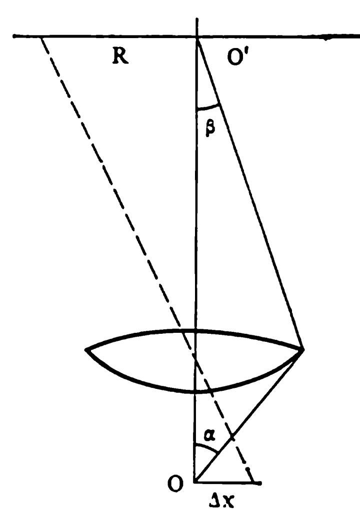


Fig. 41

Troisième exemple. Soit à déterminer la vitesse d'un corps macroscopique mobile par mesure du déplacement, par effet Doppler, de la fréquence d'une lumière monochromatique réfléchie par ce corps.

Prenons pour corps un miroir plan parfaitement réfléchissant, animé d'un mouvement suivant la normale à sa surface; la lumière

(le photon) tombe normalement sur sa surface. Supposons que le photon incident se propage dans le sens du déplacement du miroir. Le photon réfléchi devra alors se déplacer en sens inverse. En vertu des lois de la conservation de l'énergie et de l'impulsion

$$\hbar\omega_0 + {}^1/_2 m v_0^2 = \hbar\omega + {}^1/_2 m v^2, \qquad (20.5)$$

$$\hbar\omega_0/c + m v_0 = -\hbar\omega/c + m v, \qquad (20.6)$$

où m est la masse du corps, v_0 et v sont ses vitesses avant et après réflexion du photon, ω_0 et ω les fréquences du photon incident et du photon réfléchi. En récrivant ces équations sous la forme

$$m(v^2 - v_0^2) = 2\hbar(\omega_0 - \omega), \qquad (20.7)$$

$$m(v - v_0) = (\hbar/c)(\omega + \omega_0), \qquad (20.8)$$

et en les divisant membre à membre, on trouve

$$v + v_0 = 2c\frac{\omega_0 - \omega}{\omega_0 + \omega}.$$

Comme la masse m du miroir est infiniment grande par rapport à la masse du photon, on a

$$v = v_0 = c\frac{\omega_0 - \omega}{\omega_0 + \omega}. \qquad (20.9)$$

Après mesure des fréquences ω_0 et ω, l'application de cette formule permet de calculer la vitesse v du miroir. En admettant que la fréquence ω_0 est connue exactement, l'erreur Δv commise dans la détermination de la vitesse proviendra de l'imprécision de la mesure de la fréquence ω. Pour mesurer ω avec une erreur $\Delta\omega$, il faut que la durée minimale de la mesure Δt satisfasse à la condition $\Delta\omega \cdot \Delta t \sim \sim 2\pi$. En vertu de (20.9)

$$\Delta v = -c\frac{2\omega_0 \Delta\omega}{(\omega_0 + \omega)^2} \approx -c\frac{\Delta\omega}{2\omega_0}.$$

Comme les instants où se produisent les réflexions des photons sont connus à Δt près, l'erreur commise dans le calcul de v entraînera une erreur Δx dans la détermination de la coordonnée du miroir:

$$\Delta x \sim |\Delta v \cdot \Delta t| \sim \frac{c}{2\omega_0}|\Delta\omega \cdot \Delta t| \sim \frac{\pi c}{\omega_0}.$$

Or, d'après (20.8), lors de son interaction avec le photon, le miroir subit une variation incontrôlable de son impulsion $\Delta p \sim 2\omega_0\hbar/c$. On retrouve ainsi la relation d'incertitude $\Delta x \cdot \Delta p \sim 2\pi\hbar = h$.

7. Ces exemples montrent que les mesures, en mécanique quantique, sont, par principe, différentes des mesures de la physique classique. Bien entendu, dans les deux cas, les mesures sont entachées d'erreurs, mais il était admis, en physique classique, que l'amélioration des méthodes et des techniques de mesure permettrait,

en principe, de les rendre aussi petites que l'on veut. *La physique quantique affirme, par contre, que la précision des mesures est limitée par principe. Cette limite de précision est inhérente à la nature des choses et ne peut être reculée par aucun perfectionnement des appareils et des méthodes de mesure.* Les relations de Heisenberg fixent une limite à la précision des mesures. L'interaction entre l'appareil de mesure macroscopique et la microparticule, qui dure autant que la mesure, ne peut être rendue aussi petite que l'on veut. Lorsqu'on mesure, par exemple, la coordonnée d'une particule, l'opération de mesure conduit irrémédiablement à une altération, par principe inévitable et incontrôlable, de l'état initial de la particule, et par conséquent à une incertitude sur la valeur de l'impulsion qui sera mesurée ultérieurement. Si on intervertit l'ordre des mesures de la coordonnée et de l'impulsion de la particule, on arrivera au même résultat.

8. Notons quelques conséquences de l'existence des relations d'incertitude (20.2).

On peut constater tout d'abord qu'*il n'existe pas d'état correspondant au repos absolu d'une particule.* D'autre part, en physique macroscopique, l'impulsion de la particule est définie par la formule $p = mv$. Pour trouver la vitesse v de la particule, on mesure ses coordonnées x_1 et x_2 à deux instants rapprochés t_1 et t_2. On forme le rapport $\frac{x_2 - x_1}{t_2 - t_1}$ et on réalise le passage à la limite $t_2 \to t_1$. Cette méthode ne convient pas à la mesure de la vitesse instantanée d'une microparticule. En effet, pour que le passage à la limite soit possible, il faut que x_1 et x_2 soient déterminés exactement ; or une mesure exacte de la coordonnée modifie notablement l'impulsion de la particule. Il s'ensuit que le passage à la limite ne permet de trouver la vitesse instantanée d'une particule en aucune de ses positions. On pourrait naturellement prendre un intervalle de temps $t_2 - t_1$ long et mesurer x_1 et x_2 d'une façon imprécise. Dans ces conditions, les erreurs de mesure n'affectent que faiblement les vitesses de la particule, ainsi que la valeur du rapport $\frac{x_2 - x_1}{t_2 - t_1}$. Mais par ce procédé on déterminera non pas la *vitesse réelle de la particule, mais sa vitesse moyenne dans l'intervalle $t_2 - t_1$.* On peut déterminer l'impulsion d'une particule par mesure de la différence de potentiel qu'elle aura surmonté dans un champ électrique accélérateur, ou encore en mesurant la longueur λ de l'onde de de Broglie à l'aide d'un appareil de diffraction.

En mécanique quantique, le partage de l'énergie totale $\mathcal{E}$ en énergie cinétique et en énergie potentielle perd toute signification. En effet, l'une d'elles dépend des impulsions, et l'autre des coordonnées. Or ces variables ne peuvent avoir simultanément des valeurs bien déterminées. L'énergie $\mathcal{E}$ ne doit être déterminée et mesurée qu'en

tant qu'*énergie totale, sans la partager en énergies cinétique et potentielle.* On reviendra sur cette question plus à fond dans ce qui suit.

9. La théorie classique ne disposait pas d'un paramètre caractérisant les dimensions atomiques. La relation d'incertitude permet de définir un tel paramètre. Considérons, par exemple, un atome hydrogénoïde dont le noyau porte la charge Ze. En raisonnant selon la théorie classique, on commencera par écrire l'équation exprimant la conservation de l'énergie

$$\frac{p^2}{2m} - \frac{Ze^2}{r} = \text{const.}$$

En supposant qu'initialement l'électron se trouvait à l'infini pratiquement au repos, on doit poser const = 0. On détermine ainsi p^2, ce qui permet d'écrire

$$p^2r^2 = 2mZe^2r.$$

Ecrivons la relation d'incertitude sous la forme exacte (20.4). Puisqu'on a $\Delta r < r$ et $\Delta p < p$, on en tire $p^2r^2 > {}^1/_4\hbar^2$, et par suite $2mZe^2r > {}^1/_4\hbar^2$, d'où

$$r > \frac{1}{Z}\frac{\hbar^2}{8me^2} = \frac{1}{Z}\cdot 0{,}66\cdot 10^{-9} \text{ cm.} \qquad (20.10)$$

Cette formule donne un ordre de grandeur convenable pour la dimension de l'atome. On ne doit pas attacher beaucoup d'importance à la valeur numérique du coefficient puisque la formule (20.10) ne sert qu'à des estimations. Cette estimation montre notamment que dans le champ coulombien du noyau l'électron ne peut pas tomber sur le noyau. D'autre part, la présence de l'électron à l'intérieur du noyau atomique est incompatible avec les relations d'incertitude de Heisenberg.

Pour utiliser ce même procédé permettant d'évaluer les dimensions du noyau atomique, on devrait remplacer, dans la formule (20.10), m par la masse du proton. On obtiendrait ainsi, pour le rayon du noyau, une valeur 2000 fois plus petite que la valeur (20.10). Or c'est une valeur beaucoup trop grande (la dimension du noyau est de l'ordre de 10^{-13} cm). Ce résultat montre que *les forces coulombiennes ne suffisent pas à former le noyau.* Le noyau doit être le siège de forces plus puissantes — les *forces nucléaires* qui surpassent les forces coulombiennes de deux ordres environ.

10. A côté de la relation (20.1) on établit, en théorie ondulatoire, la relation suivante:

$$\Delta t\cdot\Delta\omega \gtrsim 2\pi \qquad (20.11)$$

(cf. t. IV, § 29). Cette relation exprime le fait qu'un processus ondulatoire limité dans le temps ne peut être monochromatique. Si la durée du processus ondulatoire est Δt, la dispersion $\Delta\omega$ des fréquences des ondes qui y participent est dans le meilleur des cas

celle que donne la relation (20.11). Il en résulte que si on ne dispose que d'un temps Δt court pour observer un processus monochromatique, sa fréquence ne pourra être connue, par principe, qu'avec une erreur obéissant à la relation (20.11).

Si on fait correspondre l'énergie $\mathcal{E} = \hbar\omega$ à la fréquence ω, la formule (20.11) se transforme en

$$\Delta t \cdot \Delta \mathcal{E} \geqslant 2\pi\hbar = h. \tag{20.12}$$

C'est la *relation d'incertitude de Heisenberg pour le temps et l'énergie.*

La relation (20.12) signifie que plus le temps d'existence d'un état ou le temps dévolu pour l'observer est court, plus l'indétermination sur l'énergie de cet état est grande. Inversement, plus ce temps est long, plus grande sera la précision avec laquelle on pourra déterminer l'énergie de l'état. S'il s'agit d'un état stationnaire, cet état peut subsister indéfiniment. C'est justement pour cela que l'énergie des états stationnaires possède des valeurs parfaitement déterminées. Pour illustrer le cas opposé on considérera une particule élémentaire instable qui se désintègre en un laps de temps très court (disons, de l'ordre de 10^{-20} s). Comme il ne peut être question d'énergie déterminée d'une telle particule, il est superflu d'imposer la condition de la conservation de l'énergie dans l'étude du processus de sa désintégration.

PROBLÈMES

1. Quelle doit être l'énergie cinétique $\mathcal{E}_{cin}$ d'un accélérateur à électrons (ou à protons) pour pouvoir étudier des structures dont la dimension linéaire est $l \sim 1$ fermi (10^{-13} cm)?

R é p o n s e. $\mathcal{E}_{cin} > mc^2 \sqrt{1 + (\lambda/l)^2} - mc^2$, où $\lambda = h/(mc)$ est la longueur de l'onde de Compton de l'électron (du proton). Pour l'électron $\mathcal{E}_{cin} > > m_e c^2 \lambda_e / l \approx 720$ MeV. Pour le proton $\mathcal{E}_{cin} > m_p c^2 \sqrt{1 + (\lambda_p/l)^2} - m_p c^2 \approx \approx 600$ MeV.

2. La durée de vie moyenne d'un atome à l'état excité est $\Delta t \sim 10^{-8}$ s. Lorsque l'atome retourne à l'état normal, il émet un photon dont la longueur d'onde moyenne est $\lambda = 500$ nm. Evaluer la largeur $\Delta\lambda$ et la largeur relative $\Delta\lambda/\lambda$ de la raie spectrale émise, en admettant qu'aucune cause étrangère ne l'élargit (c'est la largeur naturelle des raies spectrales).

R é p o n s e.

$$\Delta\lambda \sim \frac{\lambda^2}{c \cdot \Delta t} \sim 10^{-4} \text{ nm},$$

$$\frac{\Delta\lambda}{\lambda} \sim \frac{\lambda}{c \cdot \Delta t} \sim 10^{-7}.$$

3. En remarquant que la mécanique classique (ou corpusculaire) est un cas limite approché de la mécanique ondulatoire, établir la condition permettant de calculer la diffusion des particules α et des électrons sur les noyaux atomiques par la méthode classique, ainsi que le fit Rutherford pour établir la formule (9.3).

S o l u t i o n. Notons a le rayon de l'atome. On peut admettre que la particule diffusée n'est soumise à l'action des forces coulombiennes du noyau que dans les limites d'une sphère de rayon a dont le centre est confondu avec le centre de l'atome. Cette idéalisation ne peut pas être utilisée pour de très grands

paramètres de choc lorsque les angles de diffusion sont tellement petits qu'on ne les utilise pas dans les études expérimentales. Soit un faisceau de particules plan-parallèle de section droite Δx, tombant sur un atome. Voyons à quelles conditions on peut étudier la propagation du faisceau dans l'atome par la méthode classique, i.e. sans tenir compte de la diffraction des particules (chose que l'on fait dans l'approximation de l'optique géométrique). Selon les données du t. IV, § 6, on peut négliger l'extension du faisceau par diffraction lorsque $l \ll \Delta x^2/\lambda$, où l est la longueur du faisceau. Dans le cas présent, l désigne le chemin que parcourt le faisceau dans le champ coulombien de l'atome. D'après l'ordre de grandeur $l \approx a$, $\Delta x \sim a$. Ainsi, la condition de validité de l'approximation de l'optique géométrique est de la forme $a \ll a^2/\lambda$ ou $\lambda \ll a$. En y substituant l'expression de la longueur de l'onde de de Broglie $\lambda = h/p = h/\sqrt{2m\mathscr{E}}$, on trouve la condition cherchée

$$\mathscr{E} \gg h^2/(2ma^2). \tag{20.13}$$

Nous n'avons pas tenu compte des forces nucléaires. Or le rayon a_{noy} du noyau qui représente le rayon d'action des forces nucléaires, est près de 10^5 fois plus petit que le rayon a_{at} de l'atome. De ce fait, le noyau est bombardé par un nombre de particules $(a_{at}/a_{noy})^2 \approx 10^{10}$ fois plus petit que celui qui bombarde l'atome. A proximité immédiate du noyau les forces coulombiennes ne comptent pas beaucoup. Ainsi notre approximation se trouve justifiée.

Pour les calculs numériques il est commode de récrire la formule (20.13) sous la forme

$$\mathscr{E} \gg \frac{(hc)^2}{2mc^2a^2}, \tag{20.14}$$

en utilisant la valeur numérique $hc = 1{,}24 \cdot 10^{-4}$ eV·cm. L'énergie mc^2 sera exprimée en eV et la longueur a en cm. Pour une particule α, $mc^2 = 3{,}75 \cdot 10^9$ eV. En posant $a = 10^{-8}$ cm, on obtient dans ce cas

$$\mathscr{E} \gg 0{,}2 \text{ eV}.$$

L'énergie des particules α est de l'ordre de 1 MeV, de sorte que la mise en œuvre de la mécanique classique est justifiée dans ce cas.

Pour l'électron $mc^2 = 0{,}511 \cdot 10^6$ eV, et par suite on doit avoir $\mathscr{E} \gg 120$ eV. Pour les électrons ayant une énergie cinétique ~ 100 eV, on ne peut plus utiliser la mécanique classique.

Ce problème peut être résolu aussi à l'aide de principe d'incertitude de Heisenberg.

CHAPITRE IV

L'ÉQUATION DE SCHRÖDINGER ET LA QUANTIFICATION

§ 21. L'équation de Schrödinger

1. L'onde plane de de Broglie

$$\Psi = Ce^{i(kr-\omega t)} \tag{21.1}$$

est une onde très particulière, correspondant au mouvement uniforme libre d'une particule dans une direction donnée avec une impulsion déterminée. Or, même dans l'espace libre et surtout dans les champs de forces, les particules peuvent exécuter d'autres mouvements qui seront décrits par des fonctions d'onde plus compliquées. Le principal problème de la mécanique ondulatoire consiste justement à trouver les fonctions d'onde et les conséquences physiques qui en découlent dans diverses conditions. Ce problème ne peut être résolu que grâce à l'équation qui fut trouvée, en 1926, par Schrödinger. C'est l'*équation fondamentale de la mécanique quantique*, mais elle n'est valable qu'en *mécanique quantique non relativiste*, i.e. dans le cas des mouvements qui sont lents par rapport à la vitesse de la lumière dans le vide.

L'équation de Schrödinger doit être une équation générale, également utilisable pour la résolution de *tous* les problèmes et pas seulement pour traiter des cas particuliers. Par conséquent, cette équation ne doit pas contenir des valeurs de paramètres (par exemple les conditions initiales, la forme concrète des champs de forces, etc.) caractérisant des modes de mouvement particuliers. L'équation peut comporter des constantes universelles, la constante de Planck par exemple, les masses et les impulsions des particules, mais sans indication de leurs valeurs numériques. Les champs de forces dans lesquels se meuvent les particules doivent être représentés dans l'équation sous une forme générale. La situation est donc la même qu'avec les équations de Newton et de Maxwell qui sont arrangées de manière à s'appliquer à tous les problèmes et pas seulement à la résolution de problèmes particuliers de mécanique ou d'électrodynamique. On doit exiger en outre que l'équation de Schrödinger soit *linéaire* et *homogène* sur Ψ. Ainsi sera assurée la validité du principe de superposition des fonctions d'onde, dont la nécessité est imposée par les interférences et la diffraction des ondes de matière (cf. § 19, 8).

2. Pour trouver l'équation de Schrödinger, notons que l'une de ses solutions pour l'espace libre doit être l'onde plane de de Broglie (21.1). Cherchons l'équation différentielle satisfaisant aux conditions énumérées ci-dessus et dont la solution serait l'onde (21.1). En différentiant (21.1) par rapport à x, il vient

$$\frac{\partial \Psi}{\partial x} = ik_x \Psi, \quad \frac{\partial^2 \Psi}{\partial x^2} = -k_x^2 \Psi.$$

En différentiant par rapport à y et z, on obtient des expressions analogues. En sommant les dérivées secondes, on trouve

$$\nabla^2 \Psi = -k^2 \Psi = -\frac{p^2}{\hbar^2} \Psi. \tag{21.2}$$

C'est bien une équation différentielle, mais ce n'est pas encore celle qu'on cherche. En effet, comme on a présupposé que la grandeur p était constante, l'équation (21.2) décrit un mouvement concret à impulsion constante donnée. Différentions maintenant (21.1) par rapport au temps avec ω constant:

$$\frac{\partial \Psi}{\partial t} = -i\omega \Psi = -i \frac{\mathscr{E}}{\hbar} \Psi. \tag{21.3}$$

Cette équation ne convient pas non plus, puisqu'elle décrit le mouvement dans l'espace libre d'une particule avec une énergie cinétique constante $\mathscr{E}$. Divisons maintenant membre à membre l'équation (21.2) par l'équation (21.3), et remarquons qu'en mécanique non relativiste $\mathscr{E} = p^2/(2m)$. On obtient alors une équation linéaire homogène

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi, \tag{21.4}$$

ne contenant aucun paramètre individuel se rapportant à un mouvement particulier. Postulons que l'équation (21.4) est valable pour tous les mouvements d'une particule dans l'espace libre. L'équation obtenue est l'*équation de Schrödinger en l'absence de champs de forces.*

Généralisons maintenant l'équation (21.4) au cas de mouvements dans des champs de forces. Limitons-nous au cas de *champs à forces conservatives* qui, tout comme en mécanique classique, sont caractérisés par une *fonction potentielle* ou par l'*énergie potentielle* $U(r)$. Les seules forces ne dérivant pas d'un potentiel, que l'on rencontre en mécanique atomique, sont les *forces magnétiques*; nous nous abstiendrons de les considérer ici. En remarquant que $\hbar/\partial t$ a la dimension de l'énergie, on en conclut que les quantités $i\hbar \frac{\partial \Psi}{\partial t}$ et $U(r)\Psi$ le sont aussi. Par conséquent, l'addition du terme $U(r)\Psi$ dans le second membre de (21.4) ne modifiera pas les dimensions de cette

équation. On peut penser que l'équation ainsi obtenue

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + U(\boldsymbol{r})\Psi \tag{21.5}$$

tient convenablement compte de l'influence qu'exerce un champ à forces conservatives sur le mouvement de la particule. Cette équation est l'*équation de Schrödinger.*

Il est bien évident que le procédé qui a été mis en œuvre pour établir l'équation de Shrödinger n'en constitue pas la démonstration. L'équation de Schrödinger est la formulation d'un nouveau principe, que l'on ne peut pas déduire logiquement des principes anciens qui ne l'impliquent pas. La seule preuve de la justesse de l'équation de Schrödinger est la vérification expérimentale de toutes les conséquences qu'on peut en tirer. L'équation de Schrödinger a subi avec succès toutes les épreuves.

3. Dans l'équation (21.5) se trouve implicitement incluse la nature dualiste onde-corpuscule de la matière. Conformément à l'interprétation de la fonction d'onde Ψ, la particule n'est pas localisée. On dit couramment que la particule est « étalée » dans l'espace avec une certaine probabilité. On pourrait penser que cette circonstance serait prise en considération dans la formulation de l'équation (21.5), autrement dit, que U devrait représenter l'énergie potentielle de la particule, compte tenu de ses différentes positions et des probabilités de chacune d'elles. En réalité, l'équation (21.5) n'en tient pas compte et la fonction $U(\boldsymbol{r})$ y est traitée comme en physique classique, i.e. comme la *fonction d'une particule localisée, plus particulièrement d'une particule ponctuelle placée dans un champ de forces.* Par exemple, dans le cas de l'atome d'hydrogène, l'énergie potentielle de l'électron dans le champ du noyau est posée égale à $U(r) = -e^2/r$, donc comme si ces deux particules étaient localisées.

4. L'équation de Schrödinger est du *premier ordre* par rapport au temps. Il s'ensuit que lorsqu'on donne la fonction d'onde Ψ dans tout l'espace à un instant donné (pris pour instant initial par exemple), la fonction Ψ se trouve univoquement déterminée dans la totalité de l'espace à tous les instants ultérieurs. On ne doit pas considérer cette proposition comme exprimant le principe de causalité en mécanique quantique, vu que cette « causalité » se rapporte à la fonction d'onde Ψ et que celle-ci est liée aux objets réellement observables par des relations probabilistes. De ce fait, tout au moins sous sa forme actuelle, la mécanique quantique est une *théorie statistique par principe.*

5. L'équation de Schrödinger est *linéaire et homogène par rapport à la fonction* Ψ, comme il le faut pour que soit respecté le principe de superposition. Sous sa forme mathématique exacte, le principe de superposition se ramène à deux assertions. La première affirme que *si* Ψ_1 *et* Ψ_2 *sont deux solutions quelconques de l'équation de Schrödinger,*

toute combinaison linéaire de ces solutions $\alpha_1\Psi_1 + \alpha_2\Psi_2$ *à coefficients constants (en général complexes)* α_1 *et* α_2 *est également solution de l'équation.* La deuxième affirme que *si les fonctions d'onde* Ψ_1 *et* Ψ_2 *décrivent deux états quelconques d'un système, leur combinaison linéaire* $\alpha_1\Psi_1 + \alpha_2\Psi_2$ *décrit également un état du même système.* Il est évident que l'état de la particule dépend non pas des coefficients α_1 et α_2 mais de leur rapport α_1/α_2. L'état ne change pas si on multiplie les deux coefficients par une même constante réelle ou complexe. Cela permet, par exemple, de normer la fonction $\Psi = \alpha_1\Psi_1 + \alpha_2\Psi_2$ (à condition que l'intégrale $\int \Psi^*\Psi\, dV$, étendue à tout l'espace, soit convergente).

6. En mécanique quantique, un rôle particulier revient aux *états stationnaires.* Ce sont des états dans lesquels tous les paramètres physiques observables ne changent pas au cours du temps. La fonction d'onde Ψ elle-même ne fait pas partie de ces paramètres observables, elle est *inobservable par principe.* Ne doivent changer dans le temps que les *grandeurs physiques observables,* pouvant être déduites de Ψ selon les règles de la mécanique quantique. Il apparaît que dans les états stationnaires

$$\Psi(r, t) = \psi(r)\, e^{-i\omega t}, \tag{21.6}$$

où la fréquence ω est constante et la fonction $\psi(r)$ ne dépend pas du temps. Ne connaissant pas encore les règles d'établissement à partir de Ψ des grandeurs observables en principe, on peut cependant s'assurer que l'une de ces grandeurs — la densité de probabilité $\rho = \Psi^*\Psi$ — correspondant à l'état (21.6), ne varie pas au cours du temps. En effet,

$$\rho = \psi^*(r)\, e^{i\omega t}\cdot\psi(r)\, e^{-i\omega t} = \psi^*(r)\,\psi(r),$$

et cette dernière quantité est bien indépendante du temps.

Pour déterminer les fonctions $\psi(r)$ pour les états stationnaires, il suffit de substituer l'expression (21.6) dans l'équation (21.5), ce qui donne

$$\hbar\omega\psi = \left[-\frac{\hbar^2}{2m}\nabla^2 + U(r)\right]\psi.$$

Par analogie avec les quanta de lumière, on admettra, par hypothèse, que la quantité $\hbar\omega$ représente l'énergie totale $\mathscr{E}$ de la particule se trouvant dans un état stationnaire. Ainsi, l'énergie de l'état stationnaire est définie par l'équation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + U(r)\right]\psi = \mathscr{E}\psi. \tag{21.7}$$

Comme cette équation ne fait pas intervenir le temps, on l'appelle *équation de Schrödinger pour les états stationnaires.* L'équation (21.5)

est appelée *équation temporelle* ou *équation de Schrödinger généralisée*. Toutes les remarques concernant la fonction potentielle $U(r)$, qui avaient été faites à l'occasion de l'équation (21.5), s'appliquent également à l'équation (21.7). La fonction $U(r)$ est déterminée selon la théorie classique, comme si la particule ne possédait aucune propriété ondulatoire. Bien entendu, l'équation de Schrödinger pour les états stationnaires satisfait au principe de superposition. Mais la superposition d'états stationnaires d'énergies différentes ne sera pas un état stationnaire.

Schrödinger démontra (cf. § 22) que l'équation (21.7) apporte une solution complète au problème de la quantification de l'énergie du système. Pour cela il suffit d'entendre par $\mathcal{E}$ l'énergie du système à l'état stationnaire, sans faire d'hypothèse sur la signification physique de la fonction d'onde $\psi(r)$. Il faut cependant que les solutions $\psi(r)$ de l'équation (21.7) satisfassent à certaines *conditions naturelles* qui doivent être vérifiées à l'infini et aux points singuliers de la fonction potentielle $U(r)$. On démontrera au paragraphe suivant que *ces solutions n'existent, en général, que pour certaines valeurs de* $\mathcal{E}$. Ce sont les valeurs privilégiées de l'énergie à l'état stationnaire. On trouve notamment pour l'atome d'hydrogène exactement les mêmes valeurs de $\mathcal{E}$ que celles de l'ancienne théorie de Bohr. Ce fut le premier succès de la mécanique ondulatoire qui détermina son essor ultérieur.

7. La conjugaison de l'équation (21.7) et du principe de superposition conduit à la *règle des fréquences de Bohr*. Remarquons d'abord que tout processus physique est caractérisé par des variations dans le temps de certaines grandeurs physiques bien réelles. Or, dans les états stationnaires, ces dernières restent constantes. Il s'ensuit que la fonction d'onde, décrivant un état dans lequel évoluent des phénomènes physiques réels, doit nécessairement être *non stationnaire*. Soit un état non stationnaire simple

$$\Psi = \Psi_1 + \Psi_2, \tag{21.8}$$

résultant de la superposition de deux états stationnaires

$$\Psi_1 = \psi_1(r)\, e^{-i\omega_1 t}, \quad \Psi_2 = \psi_2(r)\, e^{-i\omega_2 t}. \tag{21.9}$$

Calculons pour cet état la grandeur observable la plus simple de toutes — la densité de probabilité ρ. On a

$$\rho = \Psi^*\Psi = (\psi_1^*\psi_1 + \psi_2^*\psi_2) + \psi_2^*\psi_1 e^{i(\omega_2-\omega_1)t} + \psi_2\psi_1^* e^{-i(\omega_2-\omega_1)t}.$$

Tenons compte de la différence de phase pouvant exister entre ψ_1 et ψ_2. Pour ce faire, posons $\psi_1 = |\psi_1|\, e^{-i\delta_1}$, $\psi_2 = |\psi_2|\, e^{-i\delta_2}$, où δ_1 et δ_2 sont des grandeurs réelles. On obtient alors

$$\rho = (|\psi_1|^2 + |\psi_2|^2) + 2\,|\psi_1|\cdot|\psi_2| \cos[(\omega_2 - \omega_1)\,t + (\delta_2 - \delta_1)]. \tag{21.10}$$

C'est la densité de probabilité (non normalisée) de l'état complexe $\Psi = \Psi_1 + \Psi_2$. Elle contient le terme constant $(|\psi_1|^2 + |\psi_2|^2)$ et le terme interférentiel $2|\psi_1|\cdot|\psi_2|\cos(\omega_{12}t + \delta_2 - \delta_1)$ qui oscille harmoniquement avec la fréquence de Bohr

$$\omega_{12} = \omega_2 - \omega_1 = \frac{\mathscr{E}_2 - \mathscr{E}_1}{\hbar}. \tag{21.11}$$

La densité de probabilité totale ρ peut varier depuis la valeur maximale $(|\psi_1| + |\psi_2|)^2$ jusqu'à la valeur minimale $(|\psi_1| - |\psi_2|)^2$. Dans les cas extrêmes, le système quantique se trouve tout entier ou sur le niveau inférieur $\mathscr{E}_1$ (dans l'état Ψ_1) ou sur le niveau supérieur $\mathscr{E}_2$ (dans l'état Ψ_2) et passe périodiquement d'un niveau à l'autre.

Ces considérations utilisent l'analogie avec l'électrodynamique classique. Elles conduisent seulement de façon naturelle à la règle des fréquences de Bohr pour les processus d'émission (21.11), mais ne prétendent pas la justifier de façon rigoureuse. Exactement comme dans la théorie initiale de Bohr, la fréquence de Bohr ω_{12} est due aux transitions quantiques entre les niveaux énergétiques. Ces transitions donnent lieu à une émission de radiations parce que la particule porte une charge électrique. En effet, en notant e la charge de la particule, la quantité ρe peut être considérée comme représentant la densité de probabilité d'une charge électrique dans l'espace. S'il s'agissait de la densité de charge et non de la densité de probabilité, on se serait trouvé dans le cas classique d'une oscillation périodique de la charge en fonction du temps. Selon les conceptions classiques, une charge oscillante doit rayonner. Il est raisonnable de s'attendre à l'émission d'un rayonnement dans le cas quantique où la densité de charge est remplacée par sa probabilité.

Il faut cependant remarquer que le terme interférentiel figurant dans (21.10) présente le caractère d'une onde stationnaire non amortie. Pour assurer l'émission continue d'un rayonnement s'échappant du système, un apport d'énergie est nécessaire. Mais en physique classique, par exemple dans l'analyse du rayonnement émis par les oscillations non amorties du dipôle de Hertz (1857-1894), la situation est exactement la même. On calcule le champ et l'intensité du rayonnement sans se préoccuper du procédé mis en œuvre pour entretenir la constance de l'amplitude et de l'énergie des oscillations.

§ 22. L'équation de Schrödinger et la quantification

1. La quantification de l'énergie résulte de ce que des limitations naturelles sont imposées aux fonctions d'ondes qui sont des solutions de l'équation de Schrödinger (21.7). Compte tenu de ces limitations, l'équation (21.7) ne peut avoir des solutions que pour certaines valeurs du paramètre $\mathscr{E}$. La situation est semblable à celle qui existe dans le cas du problème des oscillations libres d'une corde fixée aux

deux extrémités. Du fait que les extrémités de la corde sont fixées, le mouvement vibratoire se présente sous forme d'ondes stationnaires ayant des fréquences telles que la longueur de la corde soit égale à un multiple entier de demi-longueur d'onde.

Les limitations naturelles imposées aux solutions de l'équation de Schrödinger (21.7) (sous une forme quelque peu renforcée mais pratiquement toujours réalisable) peuvent être caractérisées de la manière suivante: *la fonction d'onde* $\psi(\mathbf{r})$ *et ses dérivées spatiales premières doivent être finies, uniformes et continues, même aux points (lignes et surfaces) de discontinuité de la fonction potentielle* $U(\mathbf{r})$. (L'uniformité signifie qu'après le parcours d'un contour fermé de forme arbitraire, la fonction $\psi(\mathbf{r})$ retrouve sa valeur initiale.) Les valeurs privilégiées du paramètre $\mathcal{E}$ pour lesquelles l'équation (21.7) possède des solutions satisfaisant aux limitations ci-dessus, sont appelées *valeurs propres de la grandeur* $\mathcal{E}$ pour l'équation différentielle (21.7). Les solutions correspondantes sont appelées *fonctions propres* de la même équation. *Ce sont les valeurs propres de* $\mathcal{E}$ *que l'on prend pour valeurs possibles de l'énergie dans les états stationnaires.* Les valeurs propres de l'énergie $\mathcal{E}$ peuvent être soit *discrètes*, soit *remplir de façon continue* un intervalle d'énergie fini ou infini. Dans le premier cas, on dit que le spectre énergétique est *discret* et dans le second, que le spectre est *continu.*

Pour mieux faire comprendre ces considérations, examinons le cas particulier d'une particule exécutant un mouvement unidimensionnel dans un champ à forces conservatives. Supposons que la particule se meut suivant l'axe X et que la fonction potentielle $U(x)$ a la forme d'un « puits de potentiel » symétrique de profondeur finie (fig. 42, *a*). Posons que le maximum de $U(x)$ correspond à $x = \pm\infty$ et y prend toujours la même valeur. Prenons cette valeur pour le zéro de l'échelle de l'énergie. Cela revient à poser que $U(x)$ est partout négatif et s'annule pour $x = \pm\infty$. Dans le cas unidimensionnel, l'équation de Schrödinger (21.7) prend la forme

$$\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + (\mathcal{E} - U)\psi = 0. \tag{22.1}$$

Comme les coefficients de cette équation sont réels, elle possède des solutions réelles qui seules seront prises en considération. Notons que toute solution de (22.1) multipliée par un facteur constant qui peut être complexe est également solution de l'équation (22.1). La présence d'une quantité complexe n'affectant pas $\psi^*\psi$ ne peut exercer aucune influence sur les déductions physiques de la théorie.

2. Examinons d'abord le cas $\mathcal{E} < 0$. Selon les conceptions classiques, la particule ne peut pas pénétrer dans les parties de l'espace où $U > \mathcal{E}$ puisque la différence $\mathcal{E} - U$ est égale à l'énergie cinétique $^1/_2mv^2$, qui ne peut être négative. La particule ne peut donc se déplacer qu'entre deux points où $U = \mathcal{E}$. Arrivée à l'un de ces points, la

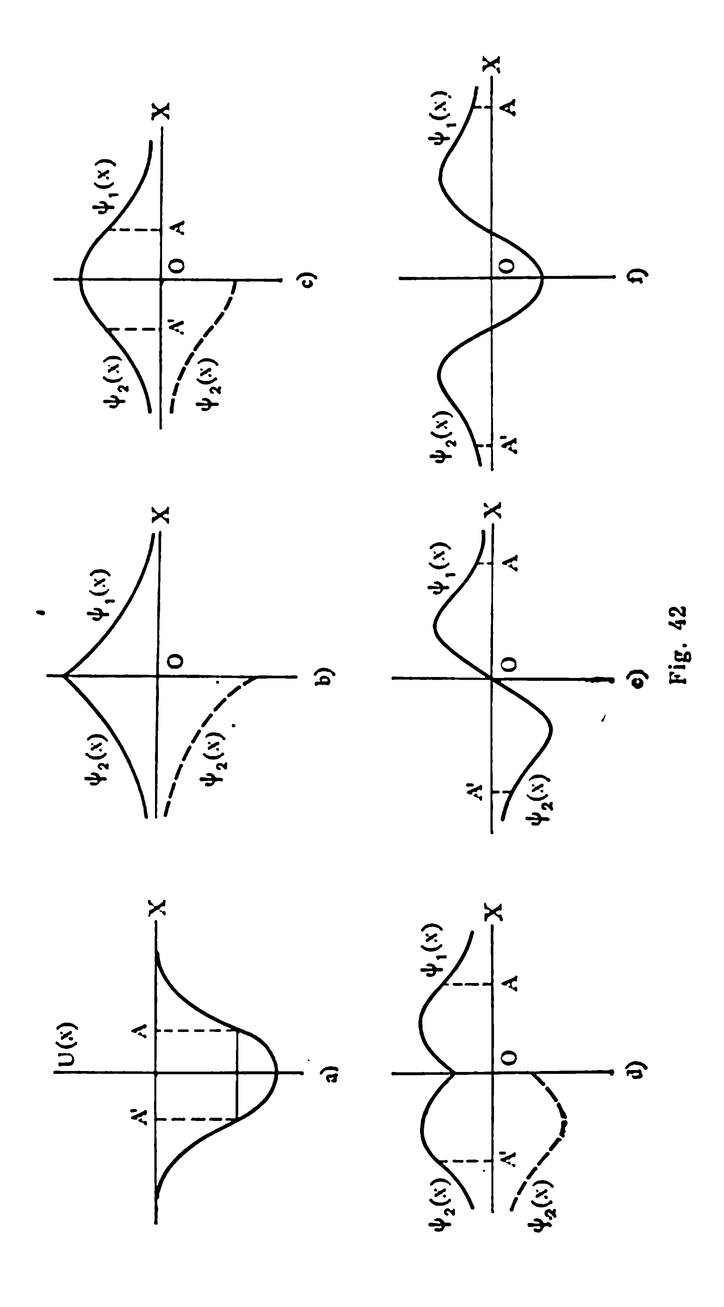


Fig. 42

particule doit rebrousser chemin. Ces points sont appelés *points de retour*. Selon la mécanique classique, tout l'espace situé au-delà des points de retour est interdit à la particule. Mais si l'équation (22.1) possède une solution qui ne soit pas identiquement nulle entre les points de retour, elle doit, au-delà de ces points, se changer en une solution qui ne soit pas identiquement nulle. En effet, comme l'équation (22.1) est du second ordre, sa solution est univoquement définie dès qu'on se fixe ψ et $d\psi/dx$ en un point de l'espace. Prenons un point situé au-delà des points de retour. Si on admet qu'en ce point $\psi \equiv 0$, conformément à ce qui vient d'être dit, la fonction ψ doit être identiquement nulle partout, y compris entre les points de retour. Il s'ensuit qu'il n'existe pas de solution qui soit différente de zéro entre les points de retour et identiquement nulle partout ailleurs.

Ainsi, la densité de probabilité $\psi^*\psi$ est différente de zéro même au-delà des points de retour. Par conséquent, selon la mécanique quantique, *il existe une probabilité finie de présence d'une particule même dans la région interdite par la mécanique classique où* $U > \mathcal{E}$. L'analyse classique donnée ci-dessus est inutilisable en mécanique quantique puisque dans ce dernier cas, en vertu du principe d'incertitude, le partage de l'énergie totale en énergie potentielle et en énergie cinétique n'a plus de sens. L'état d'une particule d'énergie $\mathcal{E}$ est décrit par une seule fonction d'onde dans tout l'intervalle infini $-\infty < x < +\infty$ et non seulement par la partie de la fonction correspondant à l'intervalle entre les points de retour classiques.

3. Pour $x \to \pm\infty$ l'équation (22.1) passe asymptotiquement à la forme

$$\frac{d^2\psi}{dx^2} - \alpha^2\psi = 0, \tag{22.2}$$

où, pour $\mathcal{E} < 0$, α est une constante positive égale à $\alpha = \sqrt{-2m\mathcal{E}/\hbar^2}$. Cette équation est vérifiée par les fonctions $C_1e^{-\alpha x}$ et $C_2e^{+\alpha x}$ avec des valeurs arbitraires des constantes C_1 et C_2. La seconde fonction devient infinie pour $x = +\infty$, tandis que la première $C_1e^{-\alpha x}$ reste régulière pour toutes les valeurs positives de x. Inversement, pour toutes les valeurs négatives de x, c'est la fonction $C_2e^{+\alpha x}$ qui reste régulière.

Examinons maintenant deux solutions de l'équation (22.1): la première $\psi_1(x)$ correspondant aux x positifs, qui, à l'infini, tend asymptotiquement vers $C_1e^{-\alpha x}$, et la seconde $\psi_2(x)$ correspondant aux x négatifs, qui, à l'infini, tend asymptotiquement vers $C_2e^{+\alpha x}$. La fonction cherchée $\psi(x)$, qui constitue la solution de (22.1) dans l'intervalle infini $-\infty < x < +\infty$, doit résulter du raccordement des solutions $\psi_1(x)$ et $\psi_2(x)$ pour $x = 0$. Ce raccordement doit assurer la continuité de la fonction $\psi(x)$ et de sa dérivée première $d\psi/dx$,

autrement dit il faut que soient vérifiées les deux égalités suivantes:

$$\psi_1(x)=\psi_2(x),\quad \frac{d\psi_1(x)}{dx}=\frac{d\psi_2(x)}{dx}\quad \text{avec}\quad x=0. \qquad (22.3)$$

Si on fixe la valeur de la constante C_1, on peut conférer à la constante C_2 une valeur telle que la première égalité soit satisfaite. La seconde égalité ne sera pas généralement vérifiée. Mais si on s'arrange pour que la deuxième égalité soit respectée, la première ne le sera pas en général. Les deux égalités ne peuvent être satisfaites simultanément que pour des valeurs déterminées de $\mathcal{E}$. Ces valeurs particulières sont les *valeurs propres de l'énergie* de la particule ou de l'équation (22.1).

Pour mettre en évidence les valeurs propres possibles de $\mathcal{E}$, on commencera par remarquer que les fonctions $d^2\psi/dx^2$ et $(E - U)\,\psi$ ont des signes contraires, comme le montre l'équation (22.1). Par conséquent, lorsque $(\mathcal{E} - U)\,\psi > 0$, la courbe $\psi = \psi(x)$ est convexe vers le haut, et lorsque $(\mathcal{E} - U)\,\psi < 0$, elle est convexe vers le bas. Aux points de retour $\mathcal{E} - U = 0$, ainsi que pour $\psi = 0$, cette courbe présente des points d'inflexion.

4. Etudions d'abord le cas où les deux points de retour sont confondus, i.e. lorsque $\mathcal{E} = U_{min}$, U_{min} étant la plus petite valeur que prend la fonction potentielle U pour $x = 0$. On a alors $\mathcal{E} - U \equiv U_{min} - U \leqslant 0$. Si on prend les fonctions $\psi_1(x)$ et $\psi_2(x)$ positives respectivement pour $x = +\infty$ et $x = -\infty$, en ces points $(\mathcal{E} - U)\,\psi_1 < 0$, $(\mathcal{E} - U)\,\psi_2 < 0$, ce qui implique qu'à l'infini les deux dérivées secondes $d^2\psi_1/dx^2$ et $d^2\psi_2/dx^2$ seront positives. Par conséquent, lorsque $|x|$ diminue, les deux courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$ seront croissantes. A mesure que $|x|$ diminue, le taux de croissance ira augmentant puisque les quantités $(\mathcal{E} - U)\,\psi_1$ et $(\mathcal{E} - U)\,\psi_2$, tout en restant négatives, croissent continuellement en module. Ainsi, lorsque $|x|$ diminue, les courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$ doivent monter constamment en restant convexes vers le bas (fig. 42, *b*). Sur la courbe $\psi = \psi_1(x)$ la dérivée première $d\psi_1/dx$ est négative, et sur la courbe $\psi = \psi_2(x)$ elle est positive. C'est pour cela qu'il est impossible de satisfaire simultanément à deux conditions (22.3). En effet, pour $x = 0$, c'est la première condition qui se trouve vérifiée, la seconde ne pouvant l'être par suite de la différence des signes des dérivées $d\psi_1/dx$ et $d\psi_2/dx$: par raccordement de $\psi_1(x)$ et $\psi_2(x)$ on voit apparaître un point anguleux sur la courbe $\psi = \psi_1 + \psi_2$ (fig. 42, *b*). On peut éliminer la discontinuité de la dérivée première en choisissant les fonctions $\psi_1(x)$ et $\psi_2(x)$ de signes contraires (sur la figure 42, *b* la fonction ψ_2 négative est représentée en pointillé). Mais dans ce cas, c'est la fonction $\psi = \psi_1 + \psi_2$ qui présentera une discontinuité au point $x = 0$. Tout cela démontre que la valeur $\mathcal{E} = U_{min}$ ne peut être une valeur propre de l'énergie.

5. Faisons croître constamment le paramètre $\mathcal{E}$; on fait alors monter la droite horizontale $U(x) = \mathcal{E}$ sur la figure 42, *a*. Les points

de retour A et A' commenceront aussitôt à s'écarter l'un de l'autre tout en restant à égales distances de l'origine des coordonnées, cela en vertu de la symétrie du puits de potentiel $U = U(x)$. Comme les points A et A' sont des points d'inflexion, entre ces points les courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$ doivent s'incurver vers le bas. Choisissons les fonctions $\psi_1(x)$ et $\psi_2(x)$ de manière à vérifier la première condition (22.3). Pour les petites valeurs de $\mathscr{E}$ les courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$ se coupent sous un certain angle pour $x = 0$. Bien entendu, du fait de la symétrie du puits de potentiel $U = U(x)$, ces courbes présentent une symétrie spéculaire par rapport à l'axe de coordonnées vertical. A mesure que le paramètre $\mathscr{E}$ augmente, i.e. à mesure que les points de retour A et A' s'éloignent l'un de l'autre, l'angle entre les courbes diminue et il arrive nécessairement que les tangentes aux courbes au point $x = 0$ deviennent horizontales. A ce moment les courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$ se raccordent insensiblement l'une à l'autre (fig. 42, *c*). La valeur correspondante du paramètre $\mathscr{E} = \mathscr{E}_1$ sera *la première* ou *la plus petite valeur propre de l'énergie*. Pour le puits de potentiel symétrique, la plus petite valeur propre $\mathscr{E}_1$ existe toujours.

6. Lors d'une augmentation ultérieure du paramètre $\mathscr{E}$ et donc d'une augmentation de la distance entre les points de retour, les courbes $\psi = \psi_1(x)$ et $\psi = \psi_2(x)$, tout en s'incurvant vers le bas, commencent à descendre. On verra apparaître à nouveau, sur l'axe vertical, un point anguleux (fig. 42, *d*), ce qui signifie que la seconde condition (22.3) cessera d'être vérifiée. Mais on peut prendre pour solution de l'équation de Schrödinger pour $x < 0$ une nouvelle fonction $\psi_2(x)$ qui se distingue de l'ancienne fonction $\psi_2(x)$ par le signe. Elle est représentée en pointillé sur la figure 42, *d*. On voit que cette courbe en pointillé est symétrique à l'ancienne courbe $\psi = \psi_2(x)$ par rapport à l'axe de coordonnées horizontal, et par conséquent symétrique à la courbe $\psi = \psi_1(x)$ par rapport à l'origine des coordonnées. Lorsque le point anguleux dont il a été question ci-dessus aura descendu jusqu'à l'origine des coordonnées, la courbe supérieure $\psi = \psi_1(x)$ se raccordera à la courbe inférieure (courbe en pointillé) et la première condition (22.3) sera satisfaite. La courbe qui en résulte est représentée en trait continu sur la figure 42, *e*. A ce moment la seconde condition (22.3) se trouvera également satisfaite et puisqu'à l'origine des coordonnées $\psi_1 = \psi_2$, elle est donc le point d'inflexion de la courbe résultante. Au cas considéré correspond la *deuxième valeur propre de l'énergie* $\mathscr{E} = \mathscr{E}_2$.

A la *troisième valeur propre* $\mathscr{E} = \mathscr{E}_3$ correspond la fonction d'onde représentée sur la figure 42, *f*.

En poursuivant ces raisonnements, on arrive à la conclusion que *les fonctions propres des états stationnaires présentent des nœuds où ces fonctions s'annulent. Le nombre de nœuds est inférieur d'une unité au numéro de la valeur propre correspondante de l'énergie.*

7. Ainsi, avec $\mathscr{E} < 0$, les valeurs propres de l'énergie forment un *spectre discret*. Comme dans ce cas la fonction d'onde de l'état stationnaire décroît asymptotiquement à l'infini suivant une loi exponentielle, on peut dire que pour $\mathscr{E} < 0$ la particule se trouve dans une région de l'espace pratiquement limitée, ce qui signifie que la particule exécute un *mouvement fini*. Rappelons qu'en mécanique classique la condition d'un mouvement fini est également de la forme $\mathscr{E} < 0$ (cf. t. I, § 25). La seule différence réside en ce qu'en mécanique classique la particule ne peut pas pénétrer dans les régions situées au-delà des points de retour, tandis que la mécanique quantique admet qu'on peut la déceler dans ces parties de l'espace, quoique avec une probabilité qui décroît rapidement à mesure qu'on s'éloigne des points de retour.

Le nombre d'états stationnaires ou de niveaux d'énergie possibles dépend de la forme de la fonction potentielle $U(x)$. Ce nombre peut être *fini* ou *infini*. Dans le cas particulier d'un puits de potentiel symétrique peu profond, il ne peut exister qu'un *seul* état stationnaire. Dans le cas où le nombre de niveaux d'énergie discrets est infiniment grand, avec l'accroissement du numéro du niveau son énergie doit évidemment tendre asymptotiquement vers $\mathscr{E} = 0$, et la distance entre les niveaux adjacents doit tendre vers zéro.

8. Notons une fois de plus qu'*il n'existe pas de niveau stationnaire d'énergie* $\mathscr{E} = U_{\min}$. S'il en existait un, la particule devrait se trouver constamment au point $x = 0$, i.e. serait au repos au fond du puits de potentiel, ce qui mettrait en défaut le principe d'incertitude de Heisenberg (cf. § 20, 8). La plus petite énergie que peut avoir une particule dans le puits de potentiel est égale à $\mathscr{E}_1$. Cette énergie est appelée *énergie au point zéro*. L'énergie au point zéro dans un puits de potentiel donné $U(x)$ ne peut être enlevée à la particule, puisque c'est la plus petite énergie possible. Pour la modifier, il faut modifier le puits de potentiel.

L'énergie au point zéro se manifeste dans de nombreux phénomènes, par exemple dans le comportement de l'hélium. Au zéro absolu de température, ses atomes ne restent pas au repos et, possédant encore de l'énergie, exécutent des *vibrations* dites *de zéro*. A proximité du zéro absolu, ces vibrations sont encore assez intenses, tandis que les forces d'attraction intermoléculaires sont faibles. Ces forces sont trop faibles pour faire passer l'hélium à l'état solide même à la température du zéro absolu. Pour le solidifier à $T = 0$, il faut que la pression soit au moins égale à 24 atm.

9. Il nous reste à trouver la solution de l'équation (22.1) pour les valeurs positives du paramètre $\mathscr{E}$. Dans ce cas à l'infini, l'équation (22.1) se réduit asymptotiquement à

$$\frac{d^2\psi}{dx^2} + \beta^2\psi = 0, \qquad (22.4)$$

où $\beta = \sqrt{2m\mathscr{E}/\hbar^2}$, ce qui signifie que β est une constante positive. Cette équation admet deux solutions linéairement indépendantes $\sin \beta x$ et $\cos \beta x$, qui restent finies pour $x = \pm\infty$. Ces deux solutions conviennent en qualité de *solutions asymptotiques* de l'équation (22.1). Notons $\psi_1(x)$ la solution de l'équation (22.1) qui, pour $x \to +\infty$, se transforme en $C_1 \cos \beta x + C_1' \sin \beta x$; d'autre part, notons $\psi_2(x)$ la solution qui, pour $x \to -\infty$, se transforme en $C_2 \cos \beta x + C_2' \sin \beta x$. Pour construire la solution générale de l'équation du second ordre (22.1), on doit trouver deux constantes arbitraires. On peut fixer arbitrairement les valeurs de deux constantes, soit C_1 et C_1', et déterminer les deux autres C_2 et C_2' à partir des conditions (22.3). Les solutions $\psi_1(x)$ et $\psi_2(x)$ pourront alors se raccorder en donnant une fonction $\psi(x)$, qui sera continue avec la dérivée première au point $x = 0$.

Par conséquent, quelle que soit la valeur du paramètre positif $\mathscr{E}$, toute solution de l'équation (22.1) peut être la fonction d'onde. Cela signifie que pour $\mathscr{E} > 0$ le spectre énergétique de la particule est *continu*. Comme pour $x = \pm\infty$ la fonction ψ reste finie, la particule peut s'éloigner à l'infini avec une probabilité différente de zéro. Autrement dit, pour $\mathscr{E} > 0$, le mouvement de la particule est *infini*. En mécanique classique, la condition de mouvement infini est exactement la même (cf. t. I, § 25).

10. Notons, pour conclure, que les conditions naturelles auxquelles doivent satisfaire les solutions de l'équation de Schrödinger et dont il a été question au début du paragraphe, peuvent être rendues moins rigoureuses. Il suffit de considérer le cas tridimensionnel. La fonction potentielle $U(r)$ peut alors devenir infinie en un certain point. Posons que ce point soit l'origine des coordonnées $r = 0$. On démontre en mécanique quantique que dans ces conditions on peut trouver des solutions qui se comportent à proximité de l'origine des coordonnées comme $\psi \sim 1/r\alpha$, α étant une constante positive, $\alpha < 1$. D'autre part, il résulte de considérations physiques que la condition d'uniformité doit concerner non pas la fonction ψ elle-même, mais la *densité de probabilité* $\psi^*\psi$. Néanmoins dans toutes les questions qu'étudie la mécanique quantique, c'est l'uniformité de la fonction ψ qui importe. Il ne nous est pas loisible d'examiner ces questions.

§ 23. L'oscillateur harmonique

1. En physique classique, on appelle *oscillateur harmonique* toute particule soumise à l'action d'une force proportionnelle au déplacement de la particule de sa position d'équilibre et orientée vers celle-ci. L'oscillateur est dit *unidimensionnel* si la particule ne peut se déplacer que suivant une droite que l'on prendra pour axe X; on fera coïncider l'origine des coordonnées avec la position d'équilibre.

La fonction potentielle de la particule est de la forme

$$U=\frac{1}{2}kx^2, \tag{23.1}$$

où k est une constante (*coefficient d'élasticité*) et x le déplacement de la particule par rapport à sa position d'équilibre. La représentation graphique de la fonction $U(x)$ est une parabole (fig. 43). Selon la mécanique classique, l'oscillateur exécute des oscillations harmoniques avec la pulsation $\omega=\sqrt{k/m}$, où m est la masse de la particule.

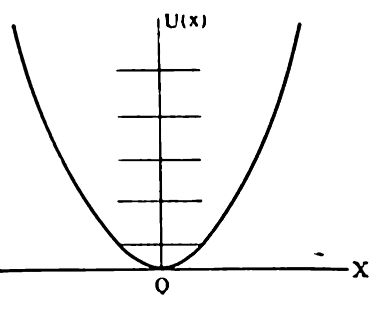


Fig. 43

En mécanique quantique, on n'utilise pas la notion de force. Aussi définit-on l'oscillateur harmonique quantique comme une particule dont la fonction potentielle est $U(x)$. Déterminons les énergies des états stationnaires de l'oscillateur en nous conformant aux considérations du paragraphe précédent. On se heurte aussitôt à une difficulté. On ne peut normaliser la fonction $U(x)$ pour qu'elle s'annule à l'infini, puisqu'elle devient elle-même infinie pour $x = \pm\infty$. Or cette difficulté est très artificielle, puisque dans les systèmes réels, lorsque $|x|$ augmente, on constate des écarts à la formule parabolique (23.1) et par suite $U(\pm\infty)$ devient fini. Considérons le cas où $U(x)$ est une fonction symétrique: $U(+\infty) = U(-\infty)$. Dans ces conditions, les méthodes du paragraphe précédent peuvent être utilisées. Il est cependant commode de prendre pour zéro de $U(x)$ sa valeur pour $x=0$. On traitera ce problème en admettant que la formule (23.1) est valable pour toutes les valeurs de x. Néanmoins les résultats obtenus ne seront applicables aux oscillateurs réels que si la valeur de $|x|$ n'est pas trop grande.

2. Pour un oscillateur harmonique unidimensionnel l'équation de Schrödinger est de la forme

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\frac{1}{2}kx^2\psi=\mathscr{E}\psi. \tag{23.2}$$

En introduisant les quantités sans dimension

$$\lambda=\frac{2\mathscr{E}}{\hbar\omega},\quad \xi=x\sqrt{\frac{k}{\hbar\omega}}, \tag{23.3}$$

l'équation (23.2) devient

$$-\frac{d^2\psi}{d\xi^2}+\xi^2\psi=\lambda\psi. \tag{23.4}$$

Pour une certaine valeur du paramètre λ cette équation a pour solution $\psi = e^{\alpha\xi^2}$, où α est une constante qu'on va définir en même temps que λ. On a

$$\frac{d\psi}{dx} = 2\alpha\xi e^{\alpha\xi^2} = 2\alpha\xi\psi,$$

$$\frac{d^2\psi}{dx^2} = 2\alpha\psi + 2\alpha\xi \frac{d\psi}{d\xi} = (4\alpha^2\xi^2 + 2\alpha)\psi.$$

En portant ces expressions dans (23.4), on obtient

$$(1 - 4\alpha^2)\xi^2 - 2\alpha = \lambda,$$

cette relation devant être identiquement vérifiée par rapport à ξ. Or cela est possible si et seulement si $1 - 4\alpha^2 = 0$, $\lambda = -2\alpha$, i.e. $\alpha = \pm 1/2$. On doit rejeter le signe $+$, car dans ce cas la fonction $\psi = e^{\alpha\xi^2}$ devient infinie pour $\xi = \pm\infty$. On arrive ainsi à la solution

$$\psi = e^{-\xi^2/2} \tag{23.5}$$

si $\lambda = 1$. Cette solution ne possède pas de nœuds et par suite elle caractérise l'*état fondamental de l'oscillateur harmonique*. A cet état correspond une énergie au point zéro égale à

$$\mathscr{E}_0 = \frac{\lambda}{2}\hbar\omega = \frac{1}{2}\hbar\omega. \tag{23.6}$$

3. Dans un état stationnaire d'énergie $\mathscr{E}_n$, la fonction ψ doit présenter n nœuds. Ce nombre de nœuds correspond à la fonction

$$\psi = P_n(\xi)\, e^{-\xi^2/2}, \tag{23.7}$$

où $P_n(\xi)$ est un polynôme de degré n ayant des racines réelles simples. Pour certaines valeurs du paramètre λ, cette fonction constitue effectivement une solution de l'équation (23.4) et s'annule à l'infini. Avec ces valeurs de λ elle représente la fonction d'onde de l'oscillateur. En la dérivant deux fois de suite et en substituant $d^2\psi/dx^2$ dans (23.4), on obtient

$$-P_n''(\xi) + \underline{2\xi P_n'(\xi)} + \underline{P_n(\xi)} = \underline{\lambda P_n(\xi)}. \tag{23.8}$$

Cette relation doit être identiquement vérifiée par rapport à ξ. Tous les termes soulignés y sont des polynômes de degré n. Le degré du polynôme $P_n''(\xi)$ est de deux unités plus petit, donc égal à $n - 2$ ($n \geqslant 2$). Pour déterminer λ, il suffit d'identifier les coefficients des termes supérieurs des polynômes soulignés. Si dans le polynôme $P_n(\xi)$ le coefficient de ξ^n est égal à a_n, dans le polynôme $2\xi P_n'(\xi)$ le coefficient correspondant est égal à $2na_n$. Il faut donc que soit vérifiée la relation $2n + 1 = \lambda$. On aura alors

$$-P_n''(\xi) + 2\xi P_n'(\xi) = 2nP_n(\xi). \tag{23.9}$$

Les polynômes qui sont les solutions de cette équation sont appelés *polynômes de Tchebychev* (1821-1894)-*Hermite*(1822-1901). On démontre que toutes les racines de ces polynômes sont simples et réelles. Il est facile de le démontrer pour n petits en déterminant les polynômes eux-mêmes et en calculant leurs racines (voir problème à la fin de ce paragraphe).

En substituant $\lambda = 2n + 1$ dans (23.3), on trouve les niveaux d'énergie de l'oscillateur

$$\mathcal{E}_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad (n = 0,\ 1,\ 2,\ \ldots). \tag{23.10}$$

Ces niveaux sont *équidistants*. Sur la figure 43 ils sont représentés par des droites horizontales.

L'oscillateur classique n'émet de la lumière que d'une *seule* fréquence ω. On pourrait penser que dans le cas quantique, conformément à la règle des fréquences de Bohr, le rayonnement émis pourrait comporter des fréquences multiples $N\omega$ (N est un nombre entier). En réalité, cela ne se produit pas dans l'émission de photons. Dans le cadre de l'ancienne théorie quantique, Bohr leva cette difficulté en s'appuyant sur le principe de correspondance. Pour éliminer les fréquences multiples on imposa une contrainte aux transitions entre les niveaux d'énergie, qui constitue la *règle de sélection*. Selon cette règle, le nombre quantique n de l'oscillateur ne peut varier que de ± 1 lors de l'émission et de l'absorption d'un photon, i.e.

$$\Delta n = \pm 1. \tag{23.11}$$

Cette règle de sélection est établie dans la mécanique quantique conséquente sans aucun recours au principe de correspondance. La mécanique quantique permet de calculer les probabilités de transition entre les niveaux d'énergie de l'oscillateur avec émission ou absorption de photons. On constate que cette probabilité est nulle chaque fois que la règle de sélection (23.11) n'est pas respectée.

PROBLÈME

Déterminer les polynômes de Tchebychev-Hermite ainsi que les fonctions d'onde d'un oscillateur harmonique unidimensionnel pour $n = 1, 2, 3, 4, 5$.

S o l u t i o n. Etudions à titre d'exemple le cas $n = 4$. Le problème consiste à trouver une solution de l'équation (23.9) sous la forme d'un polynôme

$$P_4(\xi) = a_4\xi^4 + a_3\xi^3 + a_2\xi^2 + a_1\xi + a_0.$$

En portant cette expression dans (23.9) et en identifiant les coefficients, on trouve que, conformément à la théorie générale, cette équation est vérifiée quelle que soit la valeur de a_4. On trouve ensuite $a_2 = -3a_4$, $a_0 = -{}^1/_4 a_2 = {}^3/_4 a_4$, $a_3 = a_1 = 0$. Ainsi

$$P_4(\xi) = a_4\left(\xi^4 - 3\xi^2 + \frac{3}{4}\right).$$

Les racines de ce polynôme

$$\xi = \pm \sqrt{\frac{1}{2}(3 \pm \sqrt{6})}$$

sont réelles et simples.

On trouve de façon analogue

$$P_1(\xi) = a_1\xi,$$
$$P_2(\xi) = a_2(\xi^2 - {}^1/_2),$$
$$P_3(\xi) = a_3(\xi^3 - {}^3/_2\xi),$$
$$P_5(\xi) = a_5(\xi^5 - 5\xi^3 + {}^{15}/_4\xi).$$

Pour trouver les fonctions d'onde il suffit de multiplier ces polynômes par $e^{-\xi^2/2}$. On les normalise usuellement à l'unité, i.e. on leur impose la condition

$$\int_{-\infty}^{+\infty} |P(\xi)e^{-\xi^2/2}|^2 d\xi = 1.$$

§ 24. Puits de potentiel rectangulaires unidimensionnels

1. Il est utile d'illustrer le problème de la quantification fondée sur l'équation de Schrödinger (22.1) par l'étude d'un puits de potentiel unidimensionnel symétrique et de forme rectangulaire. C'est

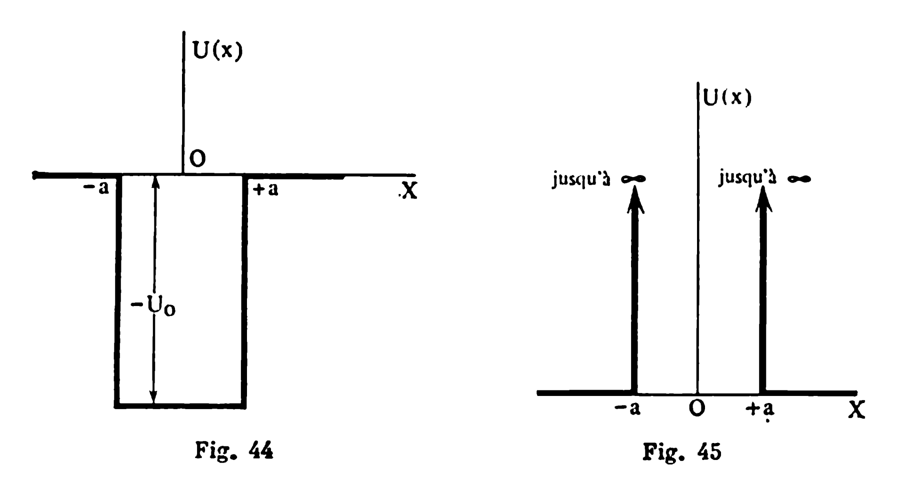


Fig. 44 Fig. 45

le nom que l'on donne à la fonction potentielle $U(x)$ qui prend une valeur constante $-U_0$ dans l'intervalle $-a < x < +a$ et une valeur nulle hors de cet intervalle (fig. 44). Il est facile de trouver pour ce cas une solution exacte de l'équation de Schrödinger, et en se fondant sur cette solution, de procéder à l'étude du problème de la quantification de l'énergie. L'intérêt des puits de potentiel rectangulaires ne se résume pas seulement à cette application. Il existe de nom-

breux cas (par exemple en physique nucléaire) où on ne connaît pas l'évolution réelle de la fonction $U(x)$. En assimilant $U(x)$ à un puits de potentiel rectangulaire, on arrive à obtenir dans ces cas une estimation qualitative et parfois même quantitative.

2. Mathématiquement, le cas le plus simple est celui d'un puits de potentiel de *profondeur infinie*, puisque alors U_0 devient infini. Il est commode de prendre pour zéro de la fonction potentielle sa valeur au fond du puits de potentiel, i.e. pour $-a < x < +a$. Sur les « parois » du puits (i.e. pour $x = \pm a$), la fonction $U(x)$ présente une discontinuité de 0 à $+\infty$. La figure 45 représente un tel puits de potentiel.

La raison de la simplification mathématique du problème résultant du passage à un puits de profondeur infinie est qu'en dehors de l'intervalle $-a < x < +a$, où U est partout infiniment grand, la fonction ψ doit s'annuler. En effet, selon la physique classique, une particule d'énergie finie $\mathcal{E}$ ne peut pénétrer dans une région de l'espace où $U(x) = +\infty$. En mécanique quantique, cette condition est remplacée par la condition d'égalité à zéro de la densité de probabilité $\psi^*\psi$, et donc de ψ. Il suffit alors d'analyser la solution de l'équation de Schrödinger dans l'intervalle $-a < x < +a$, ce qui simplifie le problème.

Dans l'intervalle $-a < x < +a$, $U(x) = 0$ et l'équation (22.1) s'écrit

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0, \tag{24.1}$$

où on a utilisé la notation

$$k^2 = \frac{2m\mathcal{E}}{\hbar}. \tag{24.2}$$

Sans restreindre la généralité, on peut ne considérer que les valeurs positives de k. La solution générale de l'équation (24.1) est de la forme

$$\psi = A \cos kx + B \sin kx,$$

sur les parois où $x = \pm a$, on doit avoir $\psi = 0$. Ainsi

$$A \cos ka + B \sin ka = 0 \quad (\text{pour } x = +a),$$

$$A \cos ka - B \sin ka = 0 \quad (\text{pour } x = -a).$$

Si $A \neq 0$, $A \cos ka = 0$ et par suite $\cos ka = 0$, $\sin ka \neq 0$, $B = 0$. Si $B \neq 0$, $B \sin ka = 0$ et donc $\sin ka = 0$, $\cos ka \neq 0$, $A = 0$. Ainsi, toutes les solutions de l'équation (24.1) peuvent être rangées dans deux catégories:

1) à fonctions paires

$$\psi = A \cos kx, \quad ka = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots,$$

2) à fonctions impaires

$\psi = B \sin kx, \quad ka = 2\frac{\pi}{2}, \ 4\frac{\pi}{2}, \ 6\frac{\pi}{2}, \ \ldots$

Dans le deuxième cas, la possibilité de ce que $ka = 0$ est exclue puisqu'on aurait alors $\psi = 0$, résultat dénué de sens physique. On détermine usuellement les constantes A et B à partir de la condition de normalisation $\int_{-a}^{+a} |\psi|^2 \, dx = 1$ (dans le cas étudié cela n'est pas nécessaire). On obtient alors

$$\psi = \begin{cases} \dfrac{1}{\sqrt{a}} \cos \dfrac{n\pi x}{2} & \text{(pour } n \text{ impairs)}, \\ \dfrac{1}{\sqrt{a}} \sin \dfrac{n\pi x}{2} & \text{(pour } n \text{ pairs)}. \end{cases} \tag{24.3}$$

Dans les deux cas, $k = n\pi/2a$, de sorte que pour tout n entier

$$\mathscr{E} = \frac{\hbar}{2m} k^2 = \frac{\hbar\pi^2}{8ma^2} n^2. \tag{24.4}$$

Ce résultat montre que l'énergie est quantifiée. Les niveaux d'énergie sont *discrets* et leur nombre devient infini pour $U_0 = +\infty$. Comme la valeur $n = 0$ est exclue, l'énergie du niveau le plus bas est égale à $\hbar\pi^2/(8ma^2)$. C'est l'énergie au point zéro dont l'existence est dictée par des considérations générales.

On peut opposer à ces résultats l'objection suivante. Sur toute surface de discontinuité de la fonction potentielle $U(x)$ on doit avoir les conditions aux limites

$$\psi_1(x-0) = \psi_2(x+0), \qquad \frac{d\psi_1(x-0)}{dx} = \frac{d\psi_2(x+0)}{dx}, \tag{24.5}$$

où $\psi_1(x)$ est la fonction $\psi(x)$ d'un côté de la surface de discontinuité, $\psi_2(x)$ la fonction $\psi(x)$ de l'autre côté de cette surface (cf. § 22). Dans notre cas, à l'intérieur de l'intervalle $-a < x < +a$ $\psi = \psi_1$ est donné par les expressions (24.3) et en dehors de cet intervalle $\psi \equiv \psi_2 = 0$. La première condition (24.5) est vérifiée et la seconde ne l'est pas. Ainsi la dérivée première de la fonction $\psi(x)$ qui vient d'être déterminée présente une discontinuité sur les parois du puits de potentiel. Néanmoins ce désaccord avec les conditions générales auxquelles doit satisfaire la fonction $\psi(x)$ n'est qu'apparent et résulte uniquement de l'*opération mathématique du passage à la limite*. Dans tous les cas réels, la profondeur U_0 du puits de potentiel est finie, quoiqu'elle puisse être grande. Dans ce dernier cas, à proximité de la paroi de part et d'autre de celle-ci, $\psi(x)$ et $d\psi/dx$ sont généralement différents de zéro et les conditions (24.5) sont rigoureusement vérifiées. Mais si on passe à la limite d'un puits de profondeur infinie,

ces conditions peuvent ne pas être satisfaites pour les valeurs limites de ces quantités. En effet, il ne résulte nullement des conditions (24.5) que les relations limites

$$\lim \psi_1 (x-0) = \lim \psi_2 (x+0),$$
$$\frac{d}{dx} \lim \psi_1 (x-0) = \frac{d}{dx} \lim \psi_2 (x+0)$$

doivent être satisfaites. C'est ce qui se produit avec les dérivées de la fonction $\psi(x)$. La solution qui a été trouvée se rapporte non pas à une fonction réelle pour un U_0 très grand, mais à la *valeur limite* de cette fonction pour $U_0 \to \infty$. Cet exemple laisse apparaître nettement l'analogie entre la question de la quantification de l'énergie et la question des vibrations d'une corde fixée aux deux extrémités. Les deux problèmes sont mathématiquement identiques si on envisage le cas d'un puits de potentiel rectangulaire infiniment profond.

3. Examinons maintenant le cas d'un puits de potentiel rectangulaire et symétrique de profondeur *finie* (fig. 44). Posons que la fonction potentielle $U(x)$ est égale à zéro hors du puits. Dans le puits $U(x) = U_0 < 0$. Plaçons l'origine des coordonnées au centre O du fond du puits. Considérons d'abord le cas où l'énergie totale $\mathscr{E}$ est négative et telle que $U_0 < \mathscr{E} < 0$. Introduisons les notations

$$k = +\sqrt{\frac{2m(\mathscr{E}-U_0)}{\hbar^2}}, \quad \alpha = +\sqrt{-\frac{2m\mathscr{E}}{\hbar^2}}. \tag{24.6}$$

A l'intérieur du puits, l'équation de Schrödinger sera alors de la forme

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0, \tag{24.7}$$

et hors du puits elle sera de la forme

$$\frac{d^2\psi}{dx^2} - \alpha^2\psi = 0. \tag{24.8}$$

La solution générale de (24.7) est de la forme

$$\psi = A \cos kx + B \sin kx. \tag{24.9}$$

L'équation (24.8) a pour solution $e^{\pm\alpha x}$. La détermination du signe doit être telle que la solution devienne égale à zéro pour $x = \pm\infty$. Ainsi hors du puits, on doit avoir

$$\psi = Ce^{-\alpha x} \quad \text{pour} \quad x > a,$$
$$\psi = De^{\alpha x} \quad \text{pour} \quad x < -a.$$

Pour raison de symétrie, la densité de probabilité $|\psi|^2$ doit être une fonction de x symétrique par rapport à l'origine des coordonnées. On doit donc avoir $C^2 = D^2$, ce qui implique deux possibilités: $C = D$ et $C = -D$. La détermination des constantes A, B, C, D

doit être telle que la fonction ψ et sa dérivée $d\psi/dx$ soient continues sur les bords du puits. Sur la frontière $x = +a$, cela conduit à

$$A \cos ka + B \sin ka = C e^{-\alpha a},$$
$$-kA \sin ka + kB \cos ka = -\alpha C e^{-\alpha a},$$

et sur la frontière $x = -a$:

$$A \cos ka - B \sin ka = D e^{-\alpha a},$$
$$kA \sin ka + kB \cos ka = \alpha D e^{-\alpha a},$$

d'où

$$2A \cos ka = (C + D)\, e^{-\alpha a},$$
$$2kA \sin ka = \alpha\, (C + D)\, e^{-\alpha a},$$
$$2B \sin ka = (C - D)\, e^{-\alpha a},$$
$$2kB \cos ka = -\alpha\, (C - D)\, e^{-\alpha a}.$$

Si $A \neq 0$ et $C = D$, on a

$$k \operatorname{tg} ka = \alpha, \qquad (24.10)$$

et si $B \neq 0$ et $C = -D$, on a

$$k \operatorname{cotg} ka = -\alpha. \qquad (24.11)$$

Ces conditions ne peuvent être satisfaites simultanément, car si elles l'étaient on aurait $k^2 = -\alpha^2$, résultat impossible puisque k et α sont tous deux réels. Lorsque tous les coefficients A, B, C, D sont nuls, la solution est dénuée de sens physique. Par conséquent, toutes les solutions possibles peuvent être rangées dans deux catégories : l'une contient les solutions à fonction d'onde paire, lorsque $A \neq 0$, $B = 0$, $C = D$, l'autre contient les solutions à fonction d'onde impaire, lorsque $A = 0$, $B \neq 0$, $C = -D$.

Pour trouver les niveaux d'énergie, on aura recours à un procédé graphique ou numérique de résolution de l'équation (24.10) ou (24.11), où les quantités positives k et α sont déterminées par les expressions (24.6). Pour appliquer la méthode graphique, on introduit les quantités sans dimension

$$\xi = ak, \quad \eta = a\alpha. \qquad (24.12)$$

On aura alors

$$\xi^2 + \eta^2 = -\frac{2mU_0 a^2}{\hbar^2}. \qquad (24.13)$$

Pour que les solutions soient à fonction d'onde paire, la relation (24.10) impose que

$$\eta = \xi \operatorname{tg} \xi, \qquad (24.10a)$$

et pour que les solutions soient à fonction d'onde impaire, on tire de (24.11)

$$\eta = -\xi \operatorname{cotg} \xi. \tag{24.11a}$$

La figure 46, *a* représente les courbes $\eta = \xi \operatorname{tg} \xi$, et la figure 46, *b* les courbes $\eta = -\xi \operatorname{cotg} \xi$. Les lignes verticales en pointillé représentent les asymptotes de ces courbes. Comme ξ et η sont positifs, on n'a à prendre en considération que les portions des courbes situées

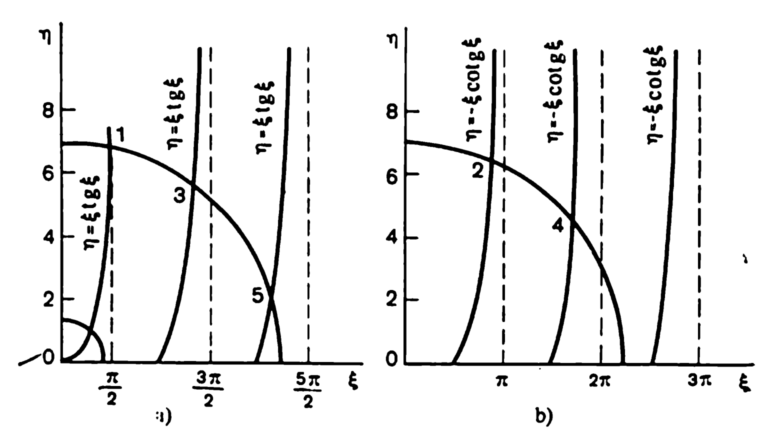


Fig. 46

dans le quadrant positif ($\xi > 0$, $\eta > 0$). Interceptons ces courbes par le cercle (24.13) dont le rayon égal à $\sqrt{-2mU_0}\, a/\hbar$ est supposé connu puisque U_0 et a sont connus. Les coordonnées des points d'intersection de ce cercle avec les courbes (24.10a) et (24.11a) indiquent les valeurs possibles de ξ et η et par suite celles de k et α. Disposant de ces données, on trouve les valeurs de $\mathscr{E}$ par les formules (24.6). Le nombre de niveaux est toujours fini et dépend de la profondeur $-U_0$ et de la largeur $2a$ du puits de potentiel. Par exemple, si le rayon du cercle est égal à 7, on obtient cinq niveaux d'énergie. Aux points d'intersection *1*, *3*, *5* correspondent des fonctions d'onde paires, et aux points *2*, *4* correspondent des fonctions d'onde impaires. Dans le cas où $0 \leqslant -U_0a^2 \leqslant \hbar^2\pi^2/(8m)$, il n'existe qu'un seul point d'intersection auquel correspond une fonction d'onde paire. Comme la quantité k est notablement différente de zéro, il résulte de (24.6) que $\mathscr{E} > U_0$. Tous les niveaux d'énergie, y compris le niveau le plus bas, se trouvent au-dessus du fond du puits de potentiel. Cette solution implique à nouveau l'existence de l'énergie au point zéro.

4. Il nous reste à examiner le cas $\mathscr{E} > 0$. La quantité α est alors un imaginaire pur: $\alpha = i\beta$. A la place de (24.8) on a l'équation

$$\frac{d^2\psi}{dx^2} + \beta^2\psi = 0 \tag{24.8a}$$

qui a pour solutions

$$\psi = A' \cos \beta x + B' \sin \beta x \quad \text{pour} \quad x > +a,$$
$$\psi = A'' \cos \beta x + B'' \sin \beta x \quad \text{pour} \quad x < -a.$$

Les deux solutions restent finies quelle que soit la valeur de x, même si elle est infiniment grande en valeur absolue. Ces solutions contiennent quatre constantes arbitraires A', B', A'', B''. Ces solutions doivent être raccordées à la solution (24.9) correspondant à $-a < x < +a$; ce raccordement doit assurer la continuité de ψ et $d\psi/dx$ sur les deux parois du puits de potentiel. On obtient ainsi quatre équations linéaires par rapport aux coefficients A', B', A'', B'' et contenant A et B en qualité de paramètres. C'est suffisant pour exprimer ces coefficients inconnus en termes de A et B, qui peuvent prendre n'importe quelles valeurs. Il s'ensuit que lorsque $\mathcal{E} > 0$ l'énergie n'est pas quantifiée et le spectre énergétique est *continu*. La fonction d'onde ne tend pas vers zéro lorsque $x \to \pm\infty$, ce qui signifie que le mouvement de la particule est *infini*, conformément à la théorie générale.

§ 25. Quantification dans le cas d'un champ de force de symétrie sphérique

1. En physique atomique, le cas où la fonction potentielle U présente une *symétrie sphérique* par rapport à un centre de force est plus important que le cas d'une fonction U unidimensionnelle. Un exemple en est le noyau atomique positivement chargé qui produit un champ électrique dans lequel se meut un électron. Nous poserons d'abord que le centre de force (noyau atomique par exemple) est *infiniment lourd* et *parfaitement fixe*. Plaçons l'origine des coordonnées au centre de force et notons $\boldsymbol{r}$ le rayon vecteur mené de l'origine des coordonnées à la particule considéreé. Si le système présente une symétrie sphérique $U = U(r)$, où $r \equiv |\boldsymbol{r}|$. Dans ce cas, la fonction d'onde ψ, i.e. la solution de l'équation de Schrödinger (21.7), peut dépendre non seulement de r, mais encore des variables angulaires caractérisant la direction du rayon vecteur $\boldsymbol{r}$. On ne considérera ici que les solutions présentant une *symétrie sphérique*, i.e. les solutions qui ne dépendent que de $|\boldsymbol{r}|$: $\psi = \psi(r)$. Dans ce cas, l'équation de Schrödinger pour les états stationnaires (21.7) se présente sous la forme

$$\frac{\hbar^2}{2m}\left(\frac{d^2\psi}{dr^2} + \frac{2}{r}\,\frac{d\psi}{dr}\right) + [\mathcal{E} - U(r)]\,\psi = 0. \qquad (25.1)$$

Cette équation diffère de l'équation unidimensionnelle (22.1) étudiée plus haut par ce qu'elle comporte le terme supplémentaire

$\frac{2}{r}\frac{d\psi}{dr}$. En procédant à la substitution

$$\psi = \frac{\chi}{r}, \tag{25.2}$$

cette dernière équation se laisse mettre sous la forme

$$\frac{\hbar^2}{2m}\frac{d^2\chi}{dr^2} + [\mathscr{E} - U(r)]\,\chi = 0. \tag{25.3}$$

Mathématiquement, cette équation est identique à l'équation (22.1) concernant le cas unidimensionnel. De ce fait, tous les résultats du § 22 restent en vigueur pour la fonction auxiliaire $\chi(r)$, à cette différence près que pour $r = 0$, la fonction χ doit être non seulement

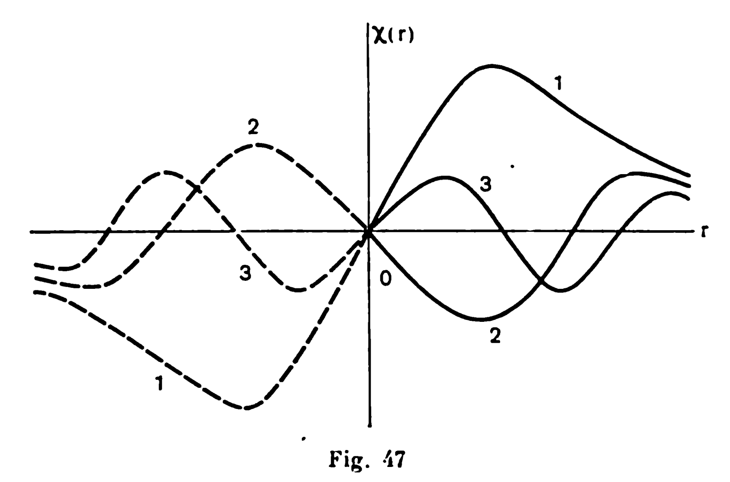


Fig. 47

finie, mais encore s'annuler. car autrement la fonction $\psi = \chi/r$ deviendrait infinie pour $r = 0$. Par conséquent la moitié des solutions qui ont été obtenues au § 22 doivent être éliminées et on ne doit garder que les solutions qui sont représentées par des courbes passant par l'origine des coordonnées (fig. 47). Pour $\mathscr{E} < 0$ ces solutions sont représentées par des courbes en trait plein pour des r positifs, et par des courbes en pointillé pour des r négatifs. Bien entendu, dans le cas étudié, les r négatifs n'ont aucune signification physique et ne sont envisagés que par formalisme mathématique pour assurer l'analogie avec le problème unidimensionnel. On peut donc utiliser ici le théorème démontré au § 22, selon lequel le nombre de nœuds de la fonction $\psi(r)$ (r est essentiellement positif) est d'une unité plus petit que le numéro de la valeur propre correspondante. Le point $r = 0$ ne doit pas être compté parmi les nœuds. Cette proposition est illustrée par la figure 47.

2. Un cas particulier d'un champ de force de symétrie sphérique est le *puits de potentiel tridimensionnel de symétrie sphérique*; la section de ce puits par un plan passant par le centre de force présente une forme rectangulaire (fig. 48). C'est le nom que l'on donne à la fonction potentielle tridimensionnelle U qui ne dépend que de la distance r jusqu'au centre de force et qui est définie par les expressions

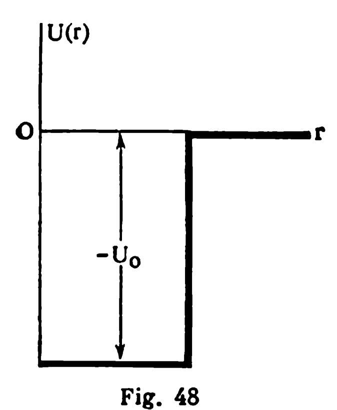


Fig. 48

$$U(r) = \begin{cases} -U_0 & \text{pour} \quad r < a, \\ 0 & \text{pour} \quad r > a. \end{cases} \tag{25.4}$$

D'après ce qui a été dit plus haut, pour trouver les fonctions d'onde $\psi(r)$ satisfaisant à la symétrie sphérique et les valeurs de l'énergie $\mathscr{E}$ des états stationnaires, on doit procéder comme dans le cas du puits de potentiel unidimensionnel rectangulaire (cf. § 24), à cette différence près que l'on doit exclure les solutions paires de l'équation (25.3) (si on prolonge les solutions dans la région des r négatifs). Cela signifie que pour $\mathscr{E} < 0$ ne subsistent que les solutions

$$\begin{aligned} \psi &= B \sin kr \quad \text{pour} \quad 0 < r < +a, \\ \psi &= Ce^{-\alpha r} \quad \text{pour} \quad r > +a, \end{aligned} \tag{25.5}$$

où k et α sont encore déterminés par les formules (24.6). En conséquence des deux formules (24.10a) et (24.11a), on ne doit garder que la seconde, i.e.

$$\eta = -\xi \operatorname{cotg} \xi, \tag{25.6}$$

ξ et η étant déterminés par les expressions (24.12).

Dans le cas d'un puits symétrique unidimensionnel, *il existe toujours au moins une valeur propre* du spectre discret d'énergie avec une fonction d'onde paire. Dans le cas d'un puits rectangulaire de symétrie sphérique, il peut arriver qu'il n'y ait aucune valeur propre. En effet, la courbe $\eta = -\xi \operatorname{cotg} \xi$ coupe l'axe de coordonnées horizontal une première fois au point $\xi = \pi/2$, $\eta = 0$ (fig. 46, *b*). D'après la formule (24.13), on voit que si

$$\left(\frac{\pi}{2}\right)^2 > -\frac{2mU_0a^2}{\hbar^2},$$

i.e.

$$-U_0 < \frac{\pi^2\hbar^2}{8ma^2}, \tag{25.7}$$

cette courbe ne coupera nulle part le cercle (24.13). Cela signifie qu'avec la condition (25.7) il ne peut exister dans le puits de potentiel

aucun niveau du spectre discret d'énergie. Le premier niveau apparaît lorsque la courbe $\eta = -\xi \operatorname{cotg} \xi$, située à l'extrême gauche sur la figure 46, *b*, commence à couper le cercle correspondant ; le point d'intersection est $\xi = \pi/2$, $\eta = 0$. D'après la seconde équation (24.6), on a alors $\mathscr{E} = 0$, ce qui signifie que si on augmente la profondeur du puits, le premier niveau apparaît à la frontière séparant les spectres discret et continu. Ce niveau se trouve à une distance du fond du puits égale à $-U_0 = \hbar^2\pi^2/(8ma^2)$; ce résultat se déduit aisément de la première formule (24.6). Dans le cas qui vient d'être considéré, le spectre discret se compose du seul niveau d'énergie au point zéro.

§ 26. Système formé par deux particules en interaction mutuelle

1. Jusqu'ici on n'a envisagé que le mouvement d'une *seule* particule dans un champ de force donné. Un cas plus proche de la réalité est celui de *deux* particules en interaction mutuelle. En mécanique classique, on dissocie le mouvement en un mouvement du système tout entier en l'absence de forces extérieures (*mouvement du centre de masse*) et un mouvement d'une particule par rapport à l'autre sous l'action des forces d'interaction s'exerçant entre elles (*mouvement relatif*). Formellement, ce dernier mouvement est ramené au mouvement d'une seule particule en remplaçant sa masse réelle par la *masse réduite*

$$m = \frac{m_1 m_2}{m_1 + m_2}, \tag{26.1}$$

où m_1 et m_2 sont les masses de la première et de la seconde particule. En mécanique quantique, on procède de la même façon pour faire l'étude des états stationnaires.

On se fonde sur l'équation de Schrödinger pour les états stationnaires de deux particules. Cette équation, qui est une généralisation évidente de (21.7), est de la forme

$$\left[-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + U(\boldsymbol{r}_1 - \boldsymbol{r}_2)\right]\psi = \mathscr{E}\psi. \tag{26.2}$$

Ici $$\nabla_1^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}, \quad \nabla_2^2 = \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2},$$

$\boldsymbol{r}_1$ (x_1, y_1, z_1) et $\boldsymbol{r}_2$ (x_2, y_2, z_2) sont les coordonnées des particules considérées. La fonction potentielle d'interaction $U(\boldsymbol{r}_1 - \boldsymbol{r}_2)$ ne dépend que de la différence des coordonnées de la première et de la seconde particule. Le plus souvent (cas des forces centrales) elle ne dépend que de leur distance de séparation $r \equiv |\boldsymbol{r}_1 - \boldsymbol{r}_2|$.

2. Transformons l'équation (26.2) pour les nouvelles variables indépendantes : les coordonnées du centre de masse

$$\boldsymbol{R}(X, Y, Z) = \frac{m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2}{m_1 + m_2} \tag{26.3}$$

et les coordonnées de la première particule par rapport à celles de la seconde

$$r\,(x,\ y,\ z) = r_1 - r_2. \tag{26.4}$$

La quantité ψ dépend maintenant de six variables: $\psi = \psi\,(r_1,\ r_2)$. Pour alléger la démonstration (sans nuire à la généralité des résultats) on ne fera les calculs que pour la fonction $\psi\,(x_1,\ x_2)$ de deux variables x_1 et x_2, et pour la fonction $\psi\,(X,\ x)$, où

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}, \quad x = x_1 - x_2.$$

En vertu de l'invariance de la différentielle totale

$$\frac{\partial\psi}{\partial x_1}\,dx_1 + \frac{\partial\psi}{\partial x_2}\,dx_2 = \frac{\partial\psi}{\partial X}\,dX + \frac{\partial\psi}{\partial x}\,dx.$$

En substituant dans le second membre, il vient

$$dX = \frac{m_1\,dx_1 + m_2\,dx_2}{m_1 + m_2}, \quad dx = dx_1 - dx_2.$$

En identifiant les coefficients, on obtient

$$\frac{\partial\psi}{\partial x_1} = \left(\frac{m_1}{m_1 + m_2}\,\frac{\partial}{\partial X} + \frac{\partial}{\partial x}\right)\psi,$$

$$\frac{\partial\psi}{\partial x_2} = \left(\frac{m_2}{m_1 + m_2}\,\frac{\partial}{\partial X} - \frac{\partial}{\partial x}\right)\psi.$$

En représentant le résultat de dérivation sous forme opérationnelle, on peut écrire aussitôt les expressions des dérivées secondes:

$$\frac{\partial^2\psi}{\partial x_1^2} = \left(\frac{m_1}{m_1 + m_2}\,\frac{\partial}{\partial X} + \frac{\partial}{\partial x}\right)^2\psi =$$

$$= \frac{m_1^2}{(m_1 + m_2)^2}\,\frac{\partial^2\psi}{\partial X^2} + 2\,\frac{m_1}{m_1 + m_2}\,\frac{\partial\psi}{\partial x} + \frac{\partial^2\psi}{\partial x^2},$$

$$\frac{\partial^2\psi}{\partial x_2^2} = \left(\frac{m_2}{m_1 + m_2}\,\frac{\partial}{\partial X} - \frac{\partial}{\partial x}\right)^2\psi =$$

$$= \frac{m_2^2}{(m_1 + m_2)^2}\,\frac{\partial^2\psi}{\partial X^2} - 2\,\frac{m_2}{m_1 + m_2}\,\frac{\partial\psi}{\partial x} + \frac{\partial^2\psi}{\partial x^2}.$$

En divisant ces relations respectivement par m_1 et m_2 et en les additionnant ensuite, il vient

$$\frac{1}{m_1}\,\frac{\partial^2\psi}{\partial x_1^2} + \frac{1}{m_2}\,\frac{\partial^2\psi}{\partial x_2^2} = \frac{1}{m_1 + m_2}\,\frac{\partial^2\psi}{\partial X^2} + \left(\frac{1}{m_1} + \frac{1}{m_2}\right)\frac{\partial^2\psi}{\partial x^2}.$$

On passe alors sans aucune difficulté aux trois variables et on écrit ensuite l'équation de Schrödinger comme suit:

$$\left[-\frac{\hbar^2}{2\,(m_1 + m_2)}\,\nabla_R^2 - \frac{\hbar^2}{2m}\,\nabla^2 + U\,(r_1 - r_2)\right]\psi = \mathcal{E}\,\psi, \tag{26.5}$$

où ∇_R^2 et ∇^2 sont les opérateurs de Laplace exprimés respectivement en fonction des variables X, Y, Z et x, y, z.

3. L'opérateur qui agit sur ψ dans le premier membre de (26.5) se décompose en une somme de deux termes indépendants, dont l'un ne dépend que de $\boldsymbol{R}$ et l'autre ne dépend que de $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. En conséquence, pour trouver la solution de l'équation (26.5) il faut résoudre les deux équations suivantes:

$$-\frac{\hbar^2}{2(m_1+m_2)}\nabla_R\psi(\boldsymbol{R}) = \mathcal{E}_R\psi(\boldsymbol{R}) \tag{26.6}$$

et

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r}) + U(\boldsymbol{r})\psi(\boldsymbol{r}) = \mathcal{E}_r\psi(\boldsymbol{r}), \tag{26.7}$$

où $\mathcal{E}_R$ et $\mathcal{E}_r$ sont des constantes vérifiant la condition $\mathcal{E}_R + \mathcal{E}_r = \mathcal{E}$. La première de ces équations décrit le *mouvement libre du centre de masse* du système, i.e. d'une particule imaginaire de masse $m_1 + m_2$. L'autre équation décrit le *mouvement relatif de la première particule par rapport à la deuxième*; la masse réelle de la particule y est remplacée par la masse réduite $m = m_1m_2/(m_1 + m_2)$. En effet, le produit $\psi(\boldsymbol{r})\psi(\boldsymbol{R})$, où $\boldsymbol{R}$ est la variable et $\boldsymbol{r}$ est considéré comme un paramètre, donc en fait comme une constante, décrit le même mouvement du centre de masse que la fonction $\psi(\boldsymbol{R})$. Le même produit $\psi(\boldsymbol{r})\psi(\boldsymbol{R})$, où on prend $\boldsymbol{r}$ pour variable et $\boldsymbol{R}$ pour paramètre, décrit le mouvement relatif. Si on multiplie (26.6) par $\psi(\boldsymbol{r})$ et (26.7) par $\psi(\boldsymbol{R})$ et si on fait la somme des deux équations membre à membre, on trouve que la fonction $\psi = \psi(\boldsymbol{r})\psi(\boldsymbol{R})$ sera la *solution générale* de l'équation (26.5). Ainsi, *la fonction d'onde commune décrivant les mouvements indépendants du centre de masse et le mouvement relatif des particules se décompose en un produit de deux fonctions de variables différentes*: $\psi(\boldsymbol{r})$ et $\psi(\boldsymbol{R})$. Tout comme en mécanique classique, l'énergie totale se décompose en une somme de l'énergie liée au mouvement du centre de masse du système et de l'énergie du mouvement relatif des particules.

Si on s'abstrait du mouvement du centre de masse, comme s'il était immobile, l'équation (26.6) peut être rejetée et ne subsiste alors que l'équation (26.7) du mouvement relatif des particules. On écrira cette équation sous la forme simple

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U(\boldsymbol{r})\psi = \mathcal{E}\psi. \tag{26.8}$$

PROBLÈMES

Le deutéron est formé d'un proton et d'un neutron liés l'un à l'autre par des *forces nucléaires* à court rayon d'action. En approximation, on peut assimiler la fonction potentielle d'interaction à un puits de potentiel spatial de forme rectangulaire de profondeur $-U_0$ et de rayon a (distance entre les centres du proton et du neutron). Le deutéron ne possède qu'un *seul* état lié. La valeur

expérimentale de son énergie de liaison est égale à 2,225 MeV. La connaissance de cette énergie ne peut suffire pour déterminer les deux inconnues U_0 et a. Posons $a = 2\cdot 10^{-13}$ cm (la valeur vraie est peu différente). On n'affirme nullement que c'est la valeur exacte de a, le seul objectif est d'expliquer le principe du procédé de calcul. Disposant de ces données, calculer la profondeur $-U_0$ du puits de potentiel.

S o l u t i o n. Seul le mouvement relatif du proton et du neutron présente de l'intérêt. On peut donc utiliser l'équation (26.8) en entendant par m la masse réduite du système. En négligeant la différence des masses du proton et du neutron, la masse réduite sera égale à $m/2$, m étant la masse de l'une des particules (du proton par exemple). Dans les calculs ultérieurs on utilise les constantes suivantes:

$$mc^2 = 938{,}28 \text{ MeV},$$

$$\hbar c = 1{,}97329\cdot 10^{-11} \text{ MeV}\cdot\text{cm}.$$

Comme le deutéron ne possède qu'un seul état lié, son énergie dans cet état est $\mathscr{E} = -2.225$ MeV. La formule (24.6) permet de trouver η, à condition d'y remplacer m par $m/2$. On obtient ainsi

$$\eta^2 = -\frac{m\mathscr{E}a^2}{\hbar^2} = -\frac{mc^2\mathscr{E}a^2}{\hbar^2c^2} = 0{,}21437;$$

$$\eta = 0{,}463099.$$

On trouve la valeur de ξ à l'aide de l'équation

$$\eta = -\xi \operatorname{cotg} \xi.$$

Pour commencer, on cherche une solution grossière par la méthode graphique en utilisant la courbe située à l'extrême gauche de la figure 46, *b*. On précise ensuite le résultat par une méthode analytique en faisant appel à l'interpolation. On trouve ainsi

$$\xi = 1{,}81993.$$

La profondeur cherchée du puits de potentiel est

$$-U_0 = \frac{\hbar^2c^2}{mc^2a^2}(\xi^2 + \eta^2) = 12{,}63 \text{ MeV}.$$

§ 27. Quantification de l'atome hydrogénoïde dans le cas d'une symétrie sphérique

1. Nous allons étudier un nouvel exemple de la quantification de l'énergie des systèmes atomiques, en l'occurrence de l'atome hydrogénoïde. Considérons le cas particulier où la fonction d'onde ψ de l'électron faisant partie de l'atome présente une *symétrie sphérique*, i.e. ne dépend que du rayon r qui est la distance entre l'électron et le noyau atomique. L'ancienne théorie de Bohr n'envisagea pas ce cas puisqu'elle ne considérait que les mouvements de l'électron autour du noyau uniquement le long d'*orbites planes*, qui ne pouvaient donc avoir une symétrie sphérique. En mécanique quantique, qui réfuta la conception du mouvement des électrons sur des orbites, il n'existe aucun obstacle à la réalisation d'états atomiques à symétrie sphérique. En raison de la symétrie sphérique, ces états seront caractérisés par une valeur nulle de la quantité qui correspond au moment de la quantité de mouvement (moment cinétique) de la

mécanique classique. Dans la théorie de Bohr, un moment cinétique nul ne pouvait être attribué qu'à un électron exécutant un mouvement rectiligne suivant le rayon. Ces mouvements auraient irrémédiablement conduit à des collisions avec le noyau atomique. Ne pouvant trouver d'autre issue à cette difficulté, l'ancienne théorie de Bohr se contentait d'exclure la possibilité des mouvements radiaux des électrons. La mécanique quantique ne connaît pas cette difficulté.

2. Soit Ze la charge du noyau. Dans le cas considéré de la symétrie sphérique, il est tout indiqué d'écrire l'équation de Schrödinger en coordonnées polaires:

$$\frac{d^2\psi}{dr^2}+\frac{2}{r}\frac{d\psi}{dr}+\left(\frac{q}{r}-\beta^2\right)\psi=0, \tag{27.1}$$

où on a utilisé les notations

$$\beta^2=-\frac{2m\mathscr{E}}{\hbar^2}, \quad q=\frac{2mZe^2}{\hbar^2}. \tag{27.2}$$

Définissons une nouvelle fonction $u(r)$ à l'aide de la formule

$$\psi=\frac{u(r)}{r}e^{-\beta r}.$$

On a alors

$$\frac{d^2u}{dr^2}-2\beta\frac{du}{dr}+\frac{q}{r}u=0. \tag{27.3}$$

Cherchons une solution de cette équation sous la forme d'une série

$$u=\sum_{k=\gamma}^{\infty}a_k r^k, \tag{27.4}$$

où γ est un nombre constant non encore déterminé. En portant (27.4) dans (27.3) et en identifiant les termes de mêmes degrés, on obtient les relations suivantes:

$$\gamma(\gamma-1)=0, \tag{27.5}$$

$$k(k+1)a_{k+1}-2\beta k a_k+q a_k=0 \quad \text{pour} \quad k\neq\gamma. \tag{27.6}$$

Il résulte de (27.5) qu'on peut avoir soit $\gamma=0$, soit $\gamma=1$. La première alternative doit être exclue puisqu'avec $\gamma=0$ le premier terme a_0 de la série (27.4) ne serait pas nul, ce qui entraînerait que pour $r=0$ la fonction ψ tendrait vers l'infini comme a_0/r; or ce résultat ne correspond pas aux conditions générales imposées à la fonction ψ aux points singuliers. Par conséquent, le développement (27.4) doit commencer par $k=1$, ce qui implique que $\gamma=1$.

Analysons le comportement de la série (27.4) à l'infini. On tire de (27.6)

$$\frac{a_{k+1}}{a_k}=\frac{2\beta k-q}{k(k+1)}. \tag{27.7}$$

De là, pour $k \to \infty$

$$\frac{a_{k+1}}{a_k} \to \frac{2\beta}{k+1}.$$

Identifions le développement (27.4) avec le développement de la fonction exponentielle:

$$e^{2\beta r} = \sum_{k=0}^{\infty} c_k r^k = \sum_{k=0}^{\infty} \frac{1}{k!} (2\beta r)^k.$$

A l'infini, les coefficients c_k de ce dernier développement se comportent de la même façon asymptotique que les coefficients a_k, puisque

$$\frac{c_{k+1}}{c_k} = \frac{2\beta}{k+1}.$$

Il s'ensuit qu'à l'infini la somme de la série (27.4) se comporte asymptotiquement comme la fonction exponentielle $e^{+2\beta r}$ et la fonction d'onde $\psi(r)$ se comporte comme $e^{\beta r}/r$; il s'ensuit que pour des valeurs arbitraires de $\mathcal{E}$ et pour $r = \infty$ la fonction $\psi(r)$ devient infinie. Néanmoins, pour les valeurs de $\mathcal{E}$ telles que la série (27.4) s'arrête, i.e. devient égale à la somme d'un nombre fini de termes, la fonction ψ ne sera pas infinie. Supposons, par exemple, que pour $k = n$ le numérateur $2\beta k - q = 0$. D'après (27.7) a_{n+1} et tous les coefficients qui suivent seront alors nuls et la série (27.4) s'arrêtera. Par conséquent, le n-ième niveau énergétique est défini par la condition $2\beta n - q = 0$. Appliquant (27.2), on trouve alors

$$\mathcal{E} = -\frac{mZ^2e^4}{2\hbar^2 n^2}, \qquad (27.8)$$

résultat qui coïncide avec la formule correspondante de la théorie de Bohr.

Ces résultats ne résolvent pas le problème de la détermination du spectre d'énergie de l'atome d'hydrogène ou des atomes hydrogénoïdes, même de la façon la plus grossière. Pour expliquer les séries spectrales, il faut associer au schéma des niveaux d'énergie les *règles de sélection* relatives à l'émission des photons. L'analyse montre que les transitions entre les niveaux d'énergie calculés, correspondant aux états de l'atome hydrogénoïde ayant une symétrie sphérique, sont des *transitions interdites*, i.e. ne donnant pas lieu à une émission (dipolaire). Pour donner une explication des séries spectrales, il est nécessaire d'envisager des états de l'atome hydrogénoïde dont la symétrie n'est pas sphérique et d'établir des règles de sélection adéquates. C'est ce qui sera fait au § 39.

3. Pour établir une comparaison avec la théorie de Bohr, déterminons encore la fonction d'onde $\psi_1(r)$ de l'état fondamental pour le cas d'une symétrie sphérique; il s'agit de l'état stationnaire de moindre énergie. Voyons quelles doivent être les valeurs des para-

mètres $\mathscr{E}$ et a_1 pour que l'équation (27.1) soit vérifiée par la fonction exponentielle

$$\psi_1(r) = e^{-\frac{r}{a_1}}, \tag{27.9}$$

avec $a_1 > 0$. Comme cette fonction n'a pas de nœuds, on peut affirmer que si $\psi_1(r)$ vérifie l'équation de Schrödinger (27.1), ce sera la fonction d'onde de l'état fondamental. En dérivant $\psi_1(r)$ deux fois de suite par rapport à r et en portant le résultat obtenu dans (27.1), on trouve

$$\frac{1}{a_1^2} - \frac{2}{a_1 r} + \frac{q}{r} - \beta^2 = 0.$$

Cette relation doit être identiquement vérifiée sur r, et par conséquent on doit avoir

$$\frac{1}{a_1^2} = \beta^2, \quad \frac{2}{a_1} = q,$$

et en vertu de (27.2)

$$a_1 = \frac{\hbar^2}{mZe^2}, \quad \mathscr{E} = \mathscr{E}_1 = -\frac{mZ^2e^4}{2\hbar^2}. \tag{27.10}$$

Cette dernière équation est un cas particulier de (27.8) pour $n = 1$, comme il convient pour l'état fondamental. Le paramètre a_1 est homogène à une longueur, et pour $Z = 1$ il représente le rayon de Bohr. Dans ce qui suit on précisera la signification physique de ce paramètre en mécanique quantique.

Pour $r = 0$, la fonction $\psi_1(r)$ devient égale à l'unité, i.e. reste finie. Par conséquent, pour $r = 0$, $u(r) = r\psi_1(r) = 0$, comme l'exige la théorie générale. La quantité $|C\psi_1|^2$, où C est un facteur de normalisation, représente la densité volumique de la probabilité de présence de l'électron dans l'espace. Introduisons encore la *densité de probabilité radiale* ρ_r. La probabilité de présence de l'électron dans une couche sphérique entre r et $r + dr$ est égale au produit du volume $4\pi r^2 dr$ de cette couche par $|C\psi_1|^2$, i.e. $4\pi C^2 r^2 |\psi_1|^2 dr$. On peut écrire cette probabilité sous la forme $\rho_r\, dr$. La quantité ρ_r est la densité de probabilité radiale puisque le produit $\rho_r\, dr$ représente la probabilité de trouver l'électron à une distance du noyau comprise entre r et $r + dr$. Par conséquent,

$$\psi_1 = Ce^{-\frac{r}{a_1}}, \quad \rho_r = 4\pi C^2 r^2 e^{-\frac{2r}{a_1}}.$$

En intégrant la seconde expression sur r entre 0 et $+\infty$ et en égalant le résultat à l'unité, on détermine la constante de normalisation C, ce qui permet d'obtenir

$$\psi_1 = \sqrt{\frac{1}{\pi a_1^3}}\, e^{-\frac{r}{a_1}}, \qquad \rho_r = \frac{4}{a_1^3} r^2 e^{-\frac{2r}{a_1}}. \tag{27.11}$$

La figure 49 représente les courbes de $|\psi_1|^2$ et de ρ_r. Les ordonnées de la première courbe ont été multipliées par 10. La courbe ρ_r passe par un maximum pour $r = a_1$. Il s'ensuit qu'en mécanique

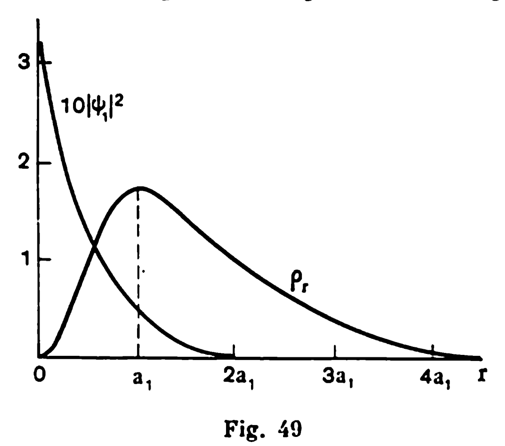


Fig. 49

quantique on doit interpréter le rayon de la première orbite de Bohr comme une distance au noyau telle que la probabilité de présence de l'électron y soit la plus grande.

PROBLÈMES

1. A quelle distance moyenne $\bar{r}$ du noyau de l'atome se trouvant dans l'état fondamental décèlera-t-on la présence de l'électron?

Réponse. $\bar{r} = {}^3/_2 a_1$.

2. Déterminer, pour le problème ci-dessus, la valeur moyenne $\overline{(1/r)}$ de l'inverse de la distance de l'électron au noyau.

Réponse. $\overline{(1/r)} = 1/a_1$.

3. Calculer les valeurs moyennes de l'énergie potentielle $\bar{U}$ et de l'énergie cinétique $\bar{\mathscr{E}}_{\text{cin}}$ de l'état fondamental de l'atome hydrogénoïde.

Réponse. $\bar{U} = -Ze^2/a_1$; $\bar{\mathscr{E}}_{\text{cin}} = Ze^2/2a_1 = -\bar{U}/2$. On notera qu'on aurait trouvé, en mécanique classique, la même relation entre U et $\mathscr{E}_{\text{cin}}$ pour un électron se mouvant sur une orbite circulaire quelconque autour du noyau.

4. Déterminer les niveaux d'énergie d'un atome hydrogénoïde se trouvant dans un état présentant une symétrie sphérique à partir du nombre de nœuds de la fonction d'onde, en procédant comme on l'avait fait au § 23 dans le cas d'un oscillateur harmonique.

Solution. Les fonctions d'onde des états excités doivent avoir des nœuds en nombre inférieur d'une unité au numéro de l'état stationnaire correspondant. Pour le n-ième état stationnaire, cette condition s'exprime par

$$\psi_n(r) = P_{n-1}(r)\, e^{-r/a_n},$$

où a_n est une constante positive et $P_{n-1}(r)$ un polynôme de degré $n-1$, dont toutes les racines sont réelles et distinctes. Il faut que la fonction $\psi_n(r)$ vérifie

l'équation de Schrödinger (27.1). On trouve par dérivation

$$\frac{d\psi_n}{dr}=\left(-\frac{1}{a_n}P_{n-1}+P'_{n-1}\right)e^{-r/a_n},$$

$$\frac{d^2\psi_n}{dr^2}=\left(\frac{1}{a_n^2}P_{n-1}-\frac{2}{a_n}P'_{n-1}+P''_{n-1}\right)e^{-r/a_n}.$$

En substituant dans (27.1), on obtient

$$\frac{1}{a_n^2}P_{n-1}-\underline{\frac{2}{a_n}P'_{n-1}}+P''_{n-1}-\underline{\frac{2}{a_n}\frac{P_{n-1}}{r}}+\frac{2P'_{n-1}}{r}+\underline{\frac{qP_{n-1}}{r}}-\beta^2P_{n-1}=0.$$

Cette relation doit être identiquement vérifiée sur r. Comme la puissance supérieure r^{n-1} ne figure que dans le premier et le dernier terme, on doit avoir

$$\frac{1}{a_n^2}=\beta^2 \quad \text{ou} \quad \frac{1}{a_n}=\beta.$$

La puissance r^{n-2} ne figure que dans les termes soulignés. Si on prend la dérivée P'_{n-1}, on verra apparaître le coefficient $(n-1)$. Par conséquent

$$-\frac{2(n-1)}{a_n}-\frac{2}{a_n}+q=0, \quad \text{ou} \quad \frac{2n}{a_n}=q.$$

On trouve ainsi

$$a_n=\frac{2n}{q}=n\frac{\hbar^2}{mZe^2}=na_1, \quad \mathscr{E}=-\frac{\hbar^2\beta}{2m}=-\frac{\hbar^2}{2ma^2}=\frac{mZ^2e^4}{2\hbar^2n^2},$$

en bon accord avec les résultats obtenus dans ce qui précède.

Le défaut de cette solution réside en ce qu'on n'a pas donné une preuve suffisante de ce que la fonction $\psi_n(r)$ est bien une solution de l'équation de Schrödinger. Pour de petites valeurs de n, il est facile de trouver la forme explicite des polynômes $P_{n-1}(r)$ ainsi que les constantes a_n correspondantes (à l'instar du calcul donné au § 23). On peut s'assurer alors que les fonctions $\psi_n(r) = P_{n-1}(r)\,e^{-r/a_n}$ vérifient bien l'équation de Schrödinger. On peut vérifier ainsi que toutes les racines du polynôme $P_{n-1}(r)$ sont bien réelles et simples.

§ 28. Les barrières de potentiel

1. Au problème de la quantification de l'énergie dans les puits de potentiel, on peut adjoindre le problème du passage des particules à travers les barrières de potentiel. On ne considérera ici que les *barrières de potentiel unidimensionnelles* où la fonction potentielle U ne dépend que d'une seule coordonnée x. Une barrière de potentiel de ce type est constituée par une région de l'espace délimitée par des plans parallèles, à l'intérieur de laquelle la fonction potentielle $U(x)$ est plus grande que dans les régions adjacentes.

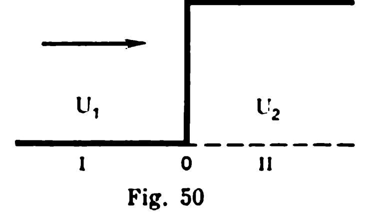


Fig. 50

Examinons d'abord le cas idéalisé simple d'une *barrière de potentiel rectangulaire* dont l'une des parois est éloignée à l'infini (fig. 50). On peut dire que c'est une barrière *échelonnée* parce que la fonction

potentielle $U(x)$ est représentée dans ce cas par une ligne échelonnée

$$U(x)=\begin{cases} U_1=\text{const dans la région } I, \text{ où } x<0, \\ U_2=\text{const dans la région } II, \text{ où } x>0, \end{cases} \quad (28.1)$$

avec $U_2 > U_1$. Une particule ou un jet de particules tombe sur la frontière de la barrière avec une vitesse constante en allant de gauche à droite. Du point de vue classique, la particule se comportera différemment suivant que son énergie totale $\mathcal{E}$ est plus grande ou plus petite que U_2. Dans le premier cas où $\mathcal{E} > U_2$, la particule, après avoir atteint la frontière de la barrière, poursuivra son mouvement sans changer de direction, mais avec une énergie cinétique moindre. Dans le second cas où $\mathcal{E} < U_2$, la particule ne peut surmonter la barrière, en sera réfléchie et rebroussera chemin avec la même énergie cinétique qu'avant.

2. La solution que donne la mécanique quantique est toute autre. Le mouvement de la particule est associé, quoique symboliquement, à la propagation d'une onde. L'équation fondamentale de la mécanique quantique — l'équation de Schrödinger — décrit, et ce de façon déterministe, la propagation des *ondes* et non pas le mouvement des particules. Pour passer du comportement des ondes au mouvement des particules, on fait appel aux *lois probabilistes*. C'est pour cette raison que le problème doit être reformulé et ensuite résolu pour les ondes à l'aide de l'équation de Schrödinger. Nous écrirons cette dernière équation sous la forme

$$\frac{d^2\psi}{dx^2}+k^2\psi=0, \quad (28.2)$$

où

$$k^2=\frac{2m}{\hbar^2}(\mathcal{E}-U), \quad (28.3)$$

U possédant des valeurs U_1 et U_2 différentes mais constantes des deux côtés de la frontière de la barrière de potentiel. Notons k_1 et k_2 les valeurs correspondantes de k.

On doit supposer maintenant que dans la région I se propage vers la barrière une onde monochromatique plane

$$\psi_1=e^{i(k_1x-\omega t)}.$$

Pour que les conditions aux limites, imposées à ψ et $d\psi/dx$, soient satisfaites sur la barrière de potentiel, il faut qu'une onde réfléchie existe dans la région I

$$\psi_1'=re^{-i(k_1x+\omega t)}$$

et qu'une onde transmise existe dans la région II

$$\psi_2=de^{i(k_2x-\omega t)}.$$

On posera que l'amplitude de l'onde incidente est égale à l'unité, ce qui ne peut affecter la généralité des résultats qui seront obtenus dans ce qui suit. Les constantes r et d sont appelées *facteur de réflexion en amplitude* et *facteur de transmission en amplitude* des ondes. Afin de les définir, remarquons que la fonction ψ ainsi que sa dérivée en x doivent être continues à la frontière de la barrière. Cela implique que pour $x = 0$ on doit avoir les relations

$$(\psi_1 + \psi_1') = \psi_2, \quad \frac{d}{dx}(\psi_1 + \psi_1') = \frac{d\psi_2}{dx},$$

ou

$$1 + r = d, \quad k_1 - k_1 r = k_2 r,$$

d'où

$$r = \frac{k_1 - k_2}{k_1 + k_2}, \quad d = \frac{2k_1}{k_1 + k_2}. \tag{28.4}$$

Ce sont les mêmes relations, mais écrites sous une forme différente que celles qui ont été établies par Fresnel (1788-1827) en optique pour l'incidence normale de la lumière sur la surface de séparation de deux milieux (cf. t. IV, §§ 65, 69). Ces relations sont valables aussi bien pour $U_2 > U_1$ (barrière de potentiel) que pour $U_2 < U_1$ (puits de potentiel).

3. La différence de principe qui existe entre la solution quantique et la solution classique réside en ce qu'en physique classique la particule est *localisée*, et ne l'est pas en mécanique quantique. En physique classique, on peut parler de l'énergie et de l'état d'une particule se trouvant en un lieu déterminé de l'espace sans se soucier de ce qui se passe partout ailleurs. En mécanique quantique, la situation est très différente. La solution que fournit la mécanique quantique est une onde, i.e. une entité concernant la *totalité de l'espace*. L'onde incidente est intimement liée à l'onde réfléchie et à l'onde transmise et on ne peut isoler l'une de ces ondes sans tenir compte des autres. L'énergie totale $\mathcal{E}$ se rapporte non pas à une onde donnée, mais à l'*état de la particule toute entière*, état défini par trois fonctions ψ_1, ψ_1', ψ_2. Une bonne compréhension de cet état de choses permet d'éviter de tirer des conclusions paradoxales à propos du passage des particules au travers des barrières de potentiel.

Notons encore que le problème de la détermination des facteurs de réflexion et de transmission en amplitude des ondes est un problème purement *déterministe* qu'on formule et qu'on résout dans le style classique en se fondant sur une formulation exacte de l'équation de Schrödinger et des conditions aux limites correspondantes. Néanmoins, ce ne sont pas les facteurs de réflexion et de transmission qui déterminent les grandeurs réelles auxquelles on a affaire dans les expériences. Dans les expériences, on mesure les *facteurs de réflexion et de transmission des particules* et non pas ceux des ondes. Les facteurs

de réflexion et de transmission pour les ondes et pour les particules sont liés les uns aux autres par des *relations probabilistes*. Les facteurs de réflexion et de transmission des particules seront définis ci-dessous. La réflexion des particules par une barrière de potentiel et leur transmission à travers celle-ci sont régies par des lois probabilistes.

4. Pour faire la comparaison des solutions classique et quantique, examinons d'abord le cas où $\mathcal{E} > U_2$. Dans ce cas, les ondes incidente, réfléchie et transmise sont des ondes *homogènes*. L'approche quantique se distingue de l'approche classique avant tout par ce que, dans le cas classique, il n'y a pas de flux de particules réfléchies. Dans le cas quantique, il apparaît nécessairement une onde réfléchie et donc une probabilité de trouver une particule allant à l'encontre du flux incident. Pour une onde homogène, on peut utiliser la notion de *densité de probabilité du flux de substance*. En effet, *un flux homogène n'est pas localisé*, puisqu'il est caractérisé par une *densité d'impulsion* déterminée, tandis que sa coordonnée est absolument indéterminée. On peut parler ainsi de la *vitesse de propagation de la probabilité* d'un tel flux, qui coïncide avec la vitesse classique: $v = p/m = \hbar k/m$. Enfin la densité de probabilité du flux d'une masse de substance est égale à $mv\psi^*\psi = \hbar k\psi^*\psi$. Dans l'onde incidente, cette quantité est égale à $\hbar k_1\psi_1^*\psi_1 = \hbar k_1$. Dans les ondes réfléchie et transmise, les densités de probabilité du flux d'une masse de substance sont respectivement égales à $|r|^2 \hbar k_1$ et $|d|^2 \hbar k_2$. Le rapport des densités de probabilité du flux de masse dans les ondes réfléchie et incidente est appelé *facteur de réflexion R de la particule*. On définit de la même façon le *facteur de transmission D de la particule*, que l'on appelle aussi *perméabilité* ou *transparence de la barrière*. On trouve pour ces grandeurs les relations suivantes:

$$R = |r|^2 = \left|\frac{k_1 - k_2}{k_1 + k_2}\right|^2, \quad D = \frac{k_2}{k_1}|d|^2 = \frac{4k_1k_2}{(k_1+k_2)^2}, \tag{28.5}$$

de sorte que $R + D = 1$, conformément à la loi de la conservation de la matière.

5. Examinons maintenant le cas où $\mathcal{E} < U_2$. Les formules (28.4) sont encore valables, de même que la première formule (28.5), puisque l'onde réfléchie est encore homogène. Mais comme la quantité k_2 est maintenant un *imaginaire pur*, l'onde qui se propage dans la deuxième région est *non homogène*. Par suite, le numérateur et le dénominateur de la première formule (28.5) seront des quantités complexes conjuguées. Par conséquent, $R = 1$, autrement dit la réflexion devient *totale*, comme en optique dans le cas analogue, mais l'onde dans le deuxième milieu ne disparaît pas pour cela. En posant $k_2 = i\alpha$, on trouve pour l'onde transmise

$$\psi_2 = \frac{2k_1}{k_1 + k_2} e^{-\alpha x} e^{-i\omega t}, \tag{28.6}$$

ce qui signifie que l'amplitude de l'onde transmise dans la région *II* s'amortit suivant une loi exponentielle à mesure qu'on s'éloigne de la frontière entre les régions *I* et *II*. La profondeur de pénétration l est définie comme la distance à laquelle la densité de probabilité du flux de substance décroît de e fois. On établit la formule

$$l = \frac{1}{2\alpha} = \frac{\lambda_2}{4\pi}, \tag{28.7}$$

où

$$\lambda_2 = \frac{2\pi\hbar}{\sqrt{2m(U_2 - \mathscr{E})}} = \frac{h}{\sqrt{2m(U_2 - \mathscr{E})}}. \tag{28.8}$$

L'onde pénètre donc dans la région *II*, quoiqu'elle subisse une réflexion totale et la probabilité de réflexion de la particule devienne égale à l'unité. Ce paradoxe est apparent et se trouve résolu exactement comme dans le cas de la réflexion totale de la lumière (cf. t. IV, § 66). La solution trouvée concerne l'*état stationnaire*, puisqu'elle est fondée sur l'équation de Schrödinger écrite pour les états stationnaires. Or, la pénétration de l'onde dans la deuxième région se produit au cours d'une période transitoire lorsque l'état ne s'est pas encore établi dans le temps. Au cours de cette période transitoire, la réflexion peut ne pas être totale. Une étude de la période transitoire peut être faite sur la base de l'équation de Schrödinger concernant les états non stationnaires (21.5).

6. Insistons une nouvelle fois sur le fait que dans l'état stationnaire trouvé ci-dessus et qu'on a caractérisé par les trois fonctions d'onde ψ_1, ψ_1', ψ_2, la particule n'est pas localisée. Elle peut se trouver en n'importe quel point de l'espace avec des probabilités différentes. Cet état spécial n'a pour caractéristique générale que le paramètre $\mathscr{E}$ que nous avons appelé *énergie totale* de la particule. Il faut faire preuve de prudence en identifiant cette notion à l'énergie totale de la mécanique classique. Il a été déjà mentionné qu'en mécanique quantique il n'est pas toujours possible de diviser l'énergie totale en énergie cinétique et en énergie potentielle (cf. § 20, 8). Pour déterminer la valeur du paramètre $\mathscr{E}$, il faut effectuer une *mesure*, i.e. exercer une certaine action sur la particule.

7. Examinons d'abord l'état de la particule dans la région *I* de l'espace. Cet état résulte de la superposition de deux ondes monochromatiques planes ψ_1 et ψ_1' se propageant l'une à l'encontre de l'autre. Leurs nombres d'onde ont des valeurs bien déterminées et égales en grandeur. Par suite les impulsions qui correspondent aux deux ondes sont égales en module. Si on mesure l'impulsion de la particule lorsqu'elle se trouve dans la région *I*, on constate qu'avec une certaine probabilité elle sera égale soit à $p_1' = \hbar k_1$, soit à $p_1' = -\hbar k_1$. Selon la relation d'incertitude de Heisenberg, la possibilité de trouver une valeur déterminée de l'impulsion dans chaque onde tient à ce que *la particule n'est pas localisée*. En effet, puisque l'incertitude

Δx sur la valeur de la coordonnée est infiniment grande, en vertu de (20.2) l'incertitude Δp sur la valeur de l'impulsion pour chaque onde est nulle. Compte tenu de la relation $p = \hbar k$, on peut récrire la formule (28.3) sous la forme

$$\mathcal{E} = \frac{p^2}{2m} + U, \tag{28.9}$$

qui montre que $\mathcal{E}$, tout comme en mécanique classique, est égal à la somme des énergies potentielle et cinétique. Cet accord avec la mécanique classique tient à ce que la fonction potentielle U_1 est constante partout dans l'espace I, ce qui implique que dans cette région de l'espace aucune force ne se manifeste. On arrivera au même résultat pour l'espace II, si seulement dans cet espace $\mathcal{E} > U_2$ et que l'onde y soit homogène.

8. Etudions le cas où $\mathcal{E} < U$ dans la région II. La probabilité de présence de la particule dans la région II fut considérée dans ce cas comme un paradoxe. La raison invoquée était la relation (28.9) qui implique qu'on doit avoir $\mathcal{E} > U$ puisque l'énergie cinétique $p^2/2m$ est essentiellement positive. Or, comme on l'avait déjà maintes fois noté, la formule (28.9) est une *relation de la mécanique classique* qui ne peut être appliquée lorsque $\mathcal{E} < U$. Dans ce cas, l'onde de de Broglie n'est pas homogène et les expressions usuelles de l'impulsion et de l'énergie cinétique perdent toute signification. Néanmoins, il est possible de déceler la particule dans la région II, puisque la probabilité de sa présence n'y est pas nulle et ne fait que décroître exponentiellement à mesure qu'on s'éloigne de la barrière dans le sens des x positifs. Déceler la présence de la particule signifie qu'on pourra indiquer les frontières de la zone où elle se trouvera étant décelée. Pratiquement, on ne peut déceler la particule qu'à l'intérieur d'une mince couche adjacente à la surface de la barrière, l'épaisseur de cette couche étant de l'ordre de la profondeur de pénétration l. La quantité l peut être prise comme incertitude sur la coordonnée après que la particule aura été détectée. En notant Δp l'incertitude sur l'impulsion, l'application de la relation (20.4) donne

$$\overline{\Delta p^2} \cdot l^2 \geqslant \hbar^2/4.$$

En y substituant l'expression de l déduite de (28.7) et (28.8), on obtient

$$\frac{\overline{\Delta p^2}}{2m} \geqslant U_2 - \mathcal{E}.$$

Ce résultat montre que pour pouvoir localiser la particule dans une couche d'épaisseur l de la région II, il faut lui communiquer une énergie cinétique au moins égale à $U_2 - \mathcal{E}$, donc une énergie cinétique positive. La particule peut acquérir cette énergie si elle est éclairée par un quantum de lumière de longueur d'onde assez courte (effet Compton). On conçoit fort bien qu'un tel procédé de localisation

modifie l'état quantique de la particule. Après interaction avec le quantum de lumière, la fonction d'onde de la particule ne sera différente de zéro qu'à l'intérieur de la couche d'épaisseur l que nous avons imposée, et sera nulle hors de cette couche.

Il n'est nullement obligatoire que la couche d'épaisseur l, où sera détectée la particule, soit adjacente à la barrière, elle peut se trouver n'importe où dans la région *II*. La probabilité de présence de la particule dépend de la position spatiale de la couche, mais l'énergie qu'il faut communiquer à la particule lors de sa localisation dans une couche d'épaisseur l dépend de l mais ne dépend pas de la position spatiale de la couche. L'épaisseur de la couche est déterminée par l'exponentielle $e^{-2\alpha l}$ et ne dépend pas de sa position.

Pour illustrer le rôle de la mesure, considérons l'exemple suivant. La particule doit être localisée dans une couche d'épaisseur l. Pour ce faire, on l'éclaire avec un pinceau de lumière se propageant le long de la couche perpendiculairement à l'axe X. Si on observe une diffusion de la lumière, cela signifiera que la particule se trouvait à l'intérieur de la couche considérée à l'instant de la diffusion. D'après l'Optique, la longueur d'onde de la lumière utilisée pour la localisation d'une particule doit être inférieure à l: $\lambda < l$. On tire de (28.7) et (28.8)

$$\lambda < \frac{h}{4\pi\sqrt{2m(U_2-\mathscr{E})}},$$

ou

$$\left(\frac{hc}{\lambda}\right)^2 \equiv (h\nu)^2 > 32\pi^2 mc^2 (U_2-\mathscr{E}).$$

On peut appliquer la mécanique non relativiste aux processus lorsque l'énergie $h\nu$ du quantum de lumière est petite par rapport à l'énergie propre mc^2 de la particule. En conséquence, en divisant le premier membre de l'inégalité précédente par la plus petite quantité $h\nu$ et le second membre par la quantité plus grande $32\pi^2 mc^2$, on obtient *a fortiori*

$$h\nu > U_2 - \mathscr{E}.$$

Il s'ensuit qu'on doit utiliser pour la localisation des quanta de lumière dont l'énergie est au moins égale à la différence entre l'énergie potentielle et l'énergie totale de la particule. Cette proposition exprime ce qu'on avait affirmé plus haut.

9. Pour conclure l'étude des barrières en forme de marche, on se propose d'établir plusieurs relations générales reliant les facteurs de réflexion et de transmission en amplitude des ondes de de Broglie à la frontière de la barrière. Si on inverse les sens de propagation de toutes les ondes de de Broglie sans que changent leurs amplitudes, l'équation de Schrödinger et les conditions aux limites correspondantes seront toujours vérifiées. Il s'ensuit que si l'état représenté

sur la figure 51, *a* est apparu, l'état représenté sur la figure 51, *b* est également possible. Sur ces figures l'onde de de Broglie est représentée par deux symboles. Le premier représente l'amplitude, et le second, le nombre d'onde de l'onde correspondante se propageant dans le sens positif ou négatif de l'axe X. Les sens de propagation sont repérés par des flèches. Sur la figure 51, *a* il n'y a qu'une seule onde

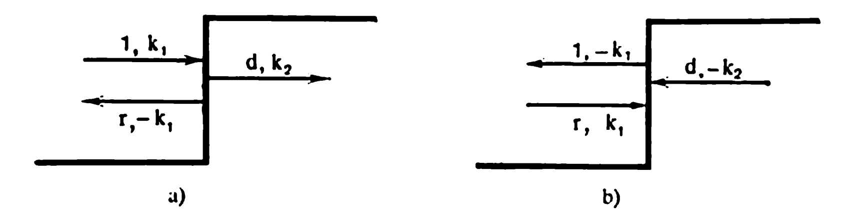


Fig. 51

incidente et sur la figure 51, *b* il y en a deux. Notons r' et d' les facteurs de réflexion et de transmission en amplitude pour une onde incidente allant de droite à gauche, i. e. de la région *II* vers la région *I*. L'onde incidente (r, k_1) donne naissance à l'onde réfléchie $(r^2, -k_1)$. L'onde incidente $(d, -k_2)$ excite l'onde transmise $(dd', -k_1)$. Les deux ondes excitées donnent par superposition l'onde s'éloignant de la barrière $(1, -k_1)$. On doit donc avoir

$$r^2 + dd' = 1. \qquad (28.10)$$

De même, l'onde (r, k_1) excite l'onde transmise (rd, k_2) et l'onde $(d, -k_2)$ excite l'onde réfléchie (dr', k_2). Les deux ondes excitées doivent se détruire mutuellement, i.e. $rd + dr' = 0$, d'où

$$r' = -r. \qquad (28.10a)$$

Les relations (28.9) et (28.10) s'appliquent également aux ondes homogènes et aux ondes non homogènes. Ces relations avaient déjà été établies dans le t. IV, § 67 pour les ondes lumineuses. Appliquons-les à une barrière de potentiel (ou à un puits de potentiel) rectangulaire de largeur finie.

La barrière avec l'onde incidente est représentée sur la figure 52, *a*. On y utilise les mêmes notations des ondes et de leurs sens de propagation que sur la figure 51. Les facteurs de réflexion et de transmission en amplitude des ondes (de gauche à droite) sur la première frontière sont dénotés par r_1 et d_1, et sur la seconde par r_2 et d_2. Pour les ondes se propageant en sens inverse (de droite à gauche), les mêmes facteurs sont dénotés par r_1', d_1', r_2', d_2'. L'image plus détaillée des ondes incidente, transmise et réfléchie est donnée sur la figure 52, *b*. Les facteurs de réflexion et de transmission en amplitude résultants pour toute la barrière sont respectivement notés r et d. Toutes les

ondes qui se forment à l'intérieur et à l'extérieur de la barrière sont représentées sur la figure 52, *b*. A l'intérieur de la barrière, deux ondes (a, k) et $(b, -k)$ se propagent en sens opposés. On peut voir sur la figure que quatre ondes tombent sur la frontière gauche de la barrière et trois sur sa frontière droite. En imposant les conditions aux limites, à savoir la continuité de ψ et $d\psi/dx$ sur chaque frontière

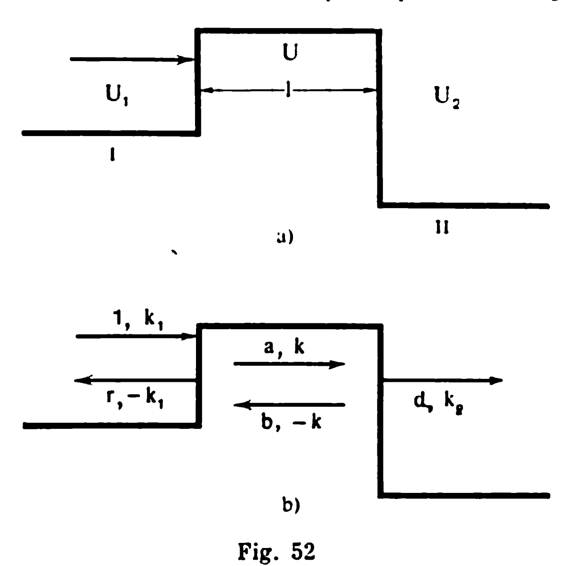


Fig. 52

de la barrière, on obtient quatre équations de premier degré à l'aide desquelles on pourra déterminer toutes les amplitudes inconnues r, a, b, d. Cependant, on peut rendre les calculs plus faciles et exprimer les résultats sous une forme plus concise en procédant de la manière suivante.

10. Examinons les conditions qui prévalent sur la frontière gauche de la barrière. Prenons cette frontière pour origine des coordonnées. Sur cette frontière viennent tomber deux ondes: $(1, k_1)$ venant de la gauche, et $(b, -k)$ venant de la droite. Ces deux ondes subissent des réflexions sur cette barrière et sont partiellement transmises au-delà. Par superposition de l'onde réfléchie et de l'onde transmise, on doit obtenir dans la première région l'onde réfléchie résultante $(r, -k_1)$ et à l'intérieur de la barrière l'onde (a, k). On doit donc avoir

$$r = r_1 + d_1'b, \quad a = d_1 + r_1'b. \tag{28.11}$$

Procédons de la même façon sur l'autre frontière de la barrière, mais en déplaçant l'origine des coordonnées sur cette frontière et en transformant en conséquence les amplitudes des ondes qui y parviennent. Les amplitudes des ondes représentées sur la figure 52, *b* sont

rapportées à l'origine des coordonnées confondue avec la frontière de gauche. Les coordonnées correspondantes sont notées x; lorsque l'origine est placée sur la frontière de droite, les coordonnées sont notées x'. Ces coordonnées sont liées par la relation $x = x' + l$. Avec le premier choix de l'origine des coordonnées, les ondes parvenant à la frontière droite de la barrière sont représentées par les expressions

$$ae^{ikx}, \quad be^{-ikx}, \quad de^{ik_2x}$$

(le facteur temporel $e^{-i\omega t}$ est omis). En remplaçant x par x', ces expressions s'écrivent

$$(ae^{ikl})\, e^{ikx'}, \quad (be^{-ikl})\, e^{-ikx'}, \quad (de^{ik_2l})\, e^{ik_2x'}.$$

Ici ce sont les expressions entre parenthèses qui assument le rôle des amplitudes des ondes. En conséquence, les conditions sur la frontière de droite sont de la forme

$$de^{ik_2l} = d_2ae^{ikl}, \quad be^{-ikl} = r_2ae^{ikl}. \tag{28.12}$$

A l'aide des équations (28.11) et (28.12) on peut déterminer toutes les inconnues r, d, a, b. Les plus intéressantes sont r et d. Compte tenu des relations $r_1' = -r_1$ et $r_1^2 + d_1d_1' = 1$, on trouve

$$\begin{aligned} r &= \frac{r_1 + r_2e^{2ikl}}{1 + r_1r_2e^{2ikl}}, \\ d &= \frac{d_1d_2e^{-i(k_2-k)l}}{1 + r_1r_2e^{2ikl}}. \end{aligned} \tag{28.13}$$

11. Disposant de ces formules, on peut calculer les facteurs de réflexion et de transmission pour les particules. Déterminons le facteur de transmission D. Supposons qu'en dehors de la barrière $U_1 < \mathcal{E}$ et $U_2 < \mathcal{E}$, et à l'intérieur de la barrière $U > \mathcal{E}$. Dans ces conditions les deux quantités

$$k_1 = \frac{1}{\hbar}\sqrt{2m(\mathcal{E} - U_1)} \qquad \text{et} \qquad k_2 = \frac{1}{\hbar}\sqrt{2m(\mathcal{E} - U_2)}$$

sont réelles et positives, tandis qu'à l'intérieur de la barrière $k = i\alpha$ avec

$$\alpha = \frac{1}{\hbar}\sqrt{2m(U - \mathcal{E})}.$$

D'autre part,

$$r_1 = \frac{k_1 - k}{k_1 + k} = \frac{k_1 - i\alpha}{k_1 + i\alpha}, \quad r_2 = \frac{k - k_2}{k + k_2} = \frac{i\alpha - k_2}{i\alpha + k_2},$$

$$d_1 = \frac{2k_1}{k_1 + k_2} = \frac{2k_1}{k_1 + i\alpha}, \quad d_2 = \frac{2k}{k + k_2} = \frac{2i\alpha}{k_2 + i\alpha}.$$

D'après (28.5) le facteur de transmission des particules se laisse calculer par la formule

$$D=\frac{k_2}{k_1}|d|^2=\frac{k_2}{k_1}dd^*. \tag{28.14}$$

Des calculs simples mais un peu longs conduisent au résultat suivant :

$$D=\frac{16k_1k_2\alpha^2}{(k_1^2+\alpha^2)(k_2^2+\alpha^2)(e^{2\alpha l}+e^{-2\alpha l})+2(\alpha^2-k_1k_2)}. \tag{28.15}$$

Dans la plupart des cas présentant de l'intérêt, on peut négliger l'exponentielle $e^{-2\alpha l}$ figurant dans le dénominateur. Supposons, par exemple, que $U-\mathscr{E}=50$ eV $=0{,}8\cdot10^{-10}$ ergs. On aura alors pour l'électron $\alpha=3{,}64\cdot10^8$ cm^{-1}, $e^{2\alpha l}=1{,}45\cdot10^3$, $e^{-2\alpha l}=0{,}69\cdot10^{-3}$. On peut négliger aussi le terme $2(\alpha^2-k_1k_2)$ qui est du même ordre de grandeur que le facteur $(k_1^2+\alpha^2)(k_2^2+\alpha^2)$. Finalement, on obtient la formule simple

$$D=D_0e^{-2\alpha l}=D_0e^{-\frac{2l}{\hbar}\sqrt{2m(U-\mathscr{E})}}, \tag{28.16}$$

où le facteur D_0 ne varie que peu lorsqu'on fait varier l, $\mathscr{E}$, U_1, U_2. On peut admettre qu'il est constant et dans nombre de cas importants le poser égal à l'unité.

Lorsqu'une particule tombe sur un des côtés de la barrière, selon les conceptions classiques, elle ne pourra pas apparaître de l'autre côté si $U>\mathscr{E}$. Par contre, la mécanique quantique accepte cette possibilité. La particule passe au travers de la région interdite $U>\mathscr{E}$ comme à travers un tunnel creusé dans la barrière. Ce phénomène fut dénommé *effet tunnel*. Dans le cas où $U_1=U_2$ (barrière de potentiel rectangulaire et symétrique), l'énergie cinétique que possède la particule derrière la barrière est la même que celle qu'elle avait en tombant sur la face avant de la barrière. La notion du passage par effet tunnel fut suggérée dès 1927 par L.I. Mandelstam (1879-1944) et M.A. Léontovitch (1903-1981). En s'appuyant sur l'équation de Schrödinger, ils ont étudié le problème de la quantification pour un oscillateur anharmonique dont la fonction potentielle était $U={}^1/_2kx^2$ pour $|x|<a$ et $U=$ const pour $|x|>a$.

12. On a examiné ci-dessus une barrière de potentiel de forme rectangulaire (fig. 52), uniquement pour simplifier l'aspect mathématique du problème. Mais, ce faisant, on n'a rien omis d'important. Pour généraliser les résultats obtenus, considérons maintenant une barrière de potentiel où U est une fonction arbitraire de x ; la figure 53 en est une illustration. La droite horizontale $U(x)=\mathscr{E}$ coupe la courbe représentant la barrière en deux points d'abscisses x_1 et x_2. En approximation, on peut remplacer la portion de la courbe située au-dessus de l'horizontale par une ligne en escalier ; ainsi toute l'aire où $\mathscr{E}<U$ sera divisée en rectangles hachurés, et chacun de

ces rectangles pourra être considéré comme une barrière de potentiel rectangulaire. Soit dx la largeur de l'un de ces rectangles et $U(x)$ sa hauteur. Si on prend dx suffisamment grand, le facteur de transmission de cette barrière rectangulaire sera donné par l'expression

$$D_0' e^{-\frac{2dx}{h}\sqrt{2m(U-\mathscr{E})}}.$$

Le facteur de transmission de la barrière toute entière s'obtient en multipliant les unes par les autres une multitude d'expressions de

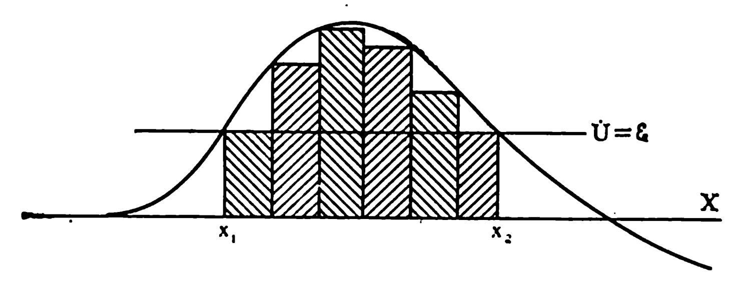


Fig. 53

cette forme. Comme les exposants s'additionnent, on obtiendra une expression de la forme

$$D = D_0 e^{-\int_{x_1}^{x_2} \frac{2}{h}\sqrt{2m(U-\mathscr{E})}\,dx}. \qquad (28.17)$$

On peut justifier cette formule approchée par un calcul mathématique rigoureux.

Le passage par effet tunnel à travers les barrières de potentiel permet d'interpréter de nombreux phénomènes physiques. C'est le cas notamment de la différence de potentiel de contact et de l'émission froide d'électrons par les métaux, faisant l'objet du paragraphe suivant. C'est encore le cas de phénomènes tels que la désintégration α, la division spontanée des noyaux atomiques ou les réactions nucléaires pour lesquelles, selon les conceptions classiques, l'énergie cinétique des noyaux entrant en collision est insuffisante pour surmonter la barrière de potentiel coulombienne, etc. Ces phénomènes seront analysés dans la partie consacrée à la physique nucléaire.

PROBLÈME

Dans une barrière (ou un puits) de potentiel rectangulaire, $U_1 = U_2$. Dans quelles conditions une particule incidente ne sera pas réfléchie par la barrière (le puits) de potentiel?

Réponse. L'énergie totale $\mathscr{E}$ doit être plus grande que l'énergie potentielle U de la particule à l'intérieur de la barrière (du puits). L'épaisseur de la barrière (du puits) doit être égale à $l = {}^1/_4\lambda, {}^3/_4\lambda, {}^5/_4\lambda, \ldots$, où $\lambda = h/\sqrt{2m\,(\mathscr{E}-U)}$ est la longueur de l'onde de de Broglie à l'intérieur de la barrière (du puits).

§ 29. Remarques sur l'interprétation de la différence de potentiel de contact. Emission froide des électrons par les métaux

1. Considérons le modèle le plus simple d'un métal — le *modèle des électrons libres*. Selon ce modèle, les électrons se comportent dans le métal comme un gaz de particules sans interactions mutuelles, se déplaçant dans un espace libre de tout champ. Les électrons sont maintenus dans le métal grâce aux forces de répulsion qui commencent à se manifester dès que les électrons s'approchent de la surface du métal. Avec ce modèle, on peut assimiler le métal à un puits de potentiel où sont enfermés les électrons. On admettra, pour simplifier, qu'il s'agit d'un puits rectangulaire d'une profondeur déterminée. Il nous est impossible de donner une justification satisfaisante de ce modèle et d'en préciser les limites de validité. Ce qui étonne surtout dans ce modèle, c'est qu'il ne tient aucun compte des forces d'interaction mutuelle des électrons, bien qu'elles soient loin d'être négligeables. Il semblerait que la possibilité de négliger l'action de ces forces réside en ce que les interactions entre les électrons ne modifient pas le nombre de niveaux énergétiques du système. Ne dépendant que du *nombre total d'électrons*, ce nombre n'est pas affecté par les forces d'interactions s'exerçant entre ceux-ci. Pour de nombreux phénomènes examinés qualitativement, c'est précisément le nombre total de niveaux énergétiques qui importe et non pas leur distribution exacte. Comme tout modèle, le modèle des électrons libres ne peut expliquer toutes les propriétés du métal, mais il fournit une explication correcte, même si elle n'est que qualitative, de nombreux effets.

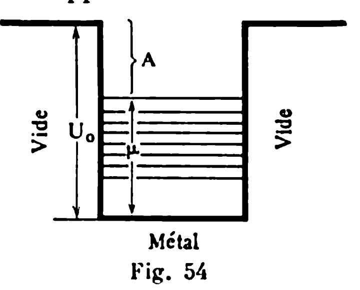


Fig. 54

La figure 54 représente le modèle du métal sous forme d'un puits de potentiel rectangulaire. A l'intérieur du métal (i.e. au fond du puits de potentiel) on pose que la fonction potentielle est nulle, et sur les parois du puits elle change par saut jusqu'à la valeur constante $U_0 > 0$. De toute évidence, à l'intérieur du puits, les niveaux énergétiques sont discrets, bien qu'ils soient très serrés dans tout échantillon macroscopique de métal. A proprement parler, on ne peut affirmer qu'il n'y ait aucune interaction entre les électrons dans le modèle des électrons libres. Cette interaction existe bien, mais il

s'agit d'interactions de type spécial, différentes des interactions de force, qui ne peuvent être expliquées dans le cadre de la mécanique classique. Il en sera question au chapitre VI. Ces interactions se manifestent en ce que chaque état quantique du système ne peut être occupé que par un seul électron. Cette proposition exprime le *principe de Pauli* (1900-1958). Dans le cas considéré, on entend par état quantique un niveau énergétique de l'électron (à condition de préciser que le nombre d'électrons pouvant occuper un même niveau

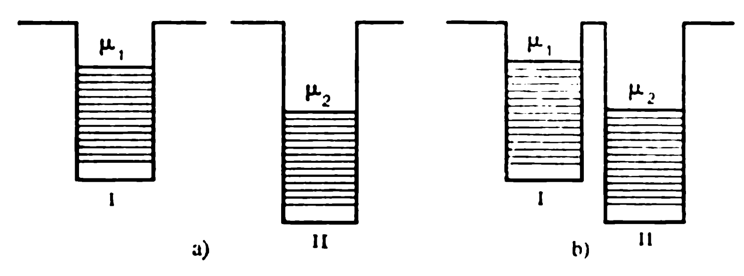


Fig. 55

d'énergie doit être multiplié par deux afin de tenir compte du spin électronique, mais cette circonstance ne joue aucun rôle dans les questions à l'étude, et nous n'en tiendrons pas compte ici).

Commençons à remplir progressivement le puits d'électrons en supposant que la température du système est égale au zéro absolu. Le premier électron introduit dans le puits viendra occuper le niveau énergétique le plus bas (niveau d'énergie zéro). Le deuxième électron occupera le niveau suivant et ainsi de suite. Le dernier électron introduit dans le puits correspond à un état qui rend le métal électriquement neutre. Ce dernier électron viendra occuper le niveau le plus élevé, noté μ, que l'on appelle *niveau* ou *énergie de Fermi* (1901-1954). Au-dessous du niveau de Fermi, tous les niveaux énergétiques du puits de potentiel sont remplis et ceux qui se trouvent au-dessus sont vacants. Rappelons qu'on avait admis que le métal se trouvait à la température égale au zéro absolu. Pour éloigner du métal l'électron occupant le niveau de Fermi, il faut fournir un travail au moins égal à

$$A = U_0 - \mu. \tag{29.1}$$

C'est le *travail de sortie* ou *d'extraction* de l'électron du métal. Bien entendu, la formule (29.1) reste valable lorsque les quantités U_0 et μ sont comptées par rapport à un niveau arbitrairement choisi au lieu de l'être à partir du fond du puits.

2. Passons maintenant à l'explication de la *différence de potentiel de contact*, qui fut découverte par Volta (1745-1827). Considérons deux métaux différents *I* et *II* (fig. 55, *a*). Convenons de compter les

positions des fonds des deux puits de potentiel et celles de tous les niveaux par rapport à un niveau commun. Le fond du puits de potentiel du métal *I* ne coïncide généralement pas avec le fond du puits du métal *II*. Il en est de même pour leurs niveaux de Fermi. Supposons, par exemple, que le niveau de Fermi du métal *I* est disposé plus haut dans l'échelle des énergies que le niveau de Fermi du métal *II*. Rapprochons les deux métaux l'un de l'autre jusqu'à ce que l'intervalle soit de l'ordre des distances atomiques, i.e. de l'ordre de 10^{-8} cm (fig. 55, *b*). Dans cet intervalle il se forme alors une barrière de potentiel très mince que les électrons peuvent traverser avec une probabilité notable et passer d'un métal à l'autre. Le passage des électrons du métal *I* dans le métal *II* se produira réellement. Mais le passage des électrons en sens inverse est irréalisable, car tous les niveaux d'énergie sur lesquels pourraient se placer dans le métal *I* les électrons venant du métal *II* sont déjà occupés. Par conséquent, le métal *I*, perdant des électrons, se chargera positivement, son potentiel ira en augmentant et son niveau de Fermi s'abaissera. Le métal *II*, par contre, acquérant des électrons, se chargera négativement, son potentiel diminuera et son niveau de Fermi s'élèvera. L'équilibre statistique s'instaurera lorsque les niveaux de Fermi des deux métaux se seront égalisés. Or, c'est justement la condition qui avait été invoquée dans le tome III (§ 104) pour expliquer l'apparition d'une différence de potentiel de contact extérieure ou intérieure. Il n'est donc nul besoin de poursuivre les explications, le lecteur étant invité à revoir le § 104 du tome III. Il importe cependant de souligner que le processus d'établissement de l'état d'équilibre se réalise par *effet tunnel* à travers la barrière de potentiel.

3. Passons à l'étude de l'émission d'électrons par les métaux. Lorsque la température du métal devient suffisamment élevée (supérieure à $\sim$1000 °C), des électrons rapides font leur apparition et ces électrons sont susceptibles de surmonter la barrière de potentiel et de s'échapper du métal. C'est l'*émission thermoionique* (cf. t. III, § 101). Or des électrons peuvent aussi être émis par un métal *froid*. Il suffit pour cela d'appliquer, suivant la normale à la surface du métal, un champ électrique fort (de l'ordre de 10^6 V/cm) dirigé vers le métal. Ce type d'émission est qualifié d'*émission froide*. L'explication de ce phénomène qui cadre *grosso modo* avec l'expérience se fonde sur la théorie du passage des électrons par effet tunnel à travers une barrière de potentiel.

Tant que le champ électrique extérieur est nul, l'énergie potentielle de l'électron est représentée sur la figure 56 par la marche *AOBC*, l'origine des coordonnées *O* se trouvant à la surface du métal. On pose que l'énergie potentielle est nulle à l'intérieur du métal et a une valeur constante *C* à l'extérieur. Si on applique un champ électrique extérieur *E* dirigé vers le métal, ce champ n'y pénétrera pas et

l'énergie potentielle de l'électron contenu dans le métal restera égale à zéro. En dehors du métal, à l'énergie potentielle C viendra s'ajouter alors l'énergie potentielle de l'électron soumis à l'action du champ électrique extérieur, qui est égale à $-eEx$ ($-e$ désigne la charge de l'électron). Cette énergie est représentée sur la figure par la droite oblique BM. Ainsi, la fonction potentielle totale de l'électron dans un champ extérieur sera représentée par les expressions

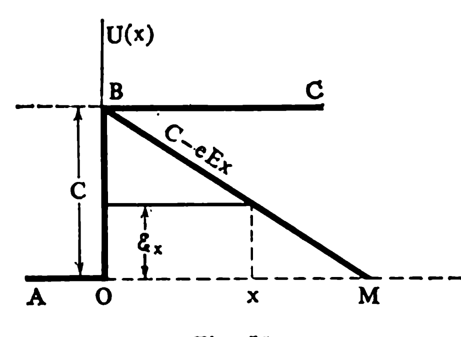


Fig. 56

$$U(x) = \begin{cases} 0 & \text{pour } x < 0, \\ C - eEx & \text{pour } x > 0. \end{cases}$$

Entre le métal et le vide s'installe une barrière de potentiel OBM. Dégageons dans le métal un groupe d'électrons d'énergie proche de $\mathcal{E}_x$. La pénétrabilité de la barrière pour des électrons de cette énergie se laisse calculer par la formule (28.17) en y posant $x_1 = 0$. Pour trouver x_2 on utilise l'équation $C - eEx_2 = \mathcal{E}_x$, d'où $x_2 = \frac{C - \mathcal{E}_x}{eE}$. Tout le problème se trouve ramené au calcul de l'intégrale

$$S = \int_0^{x_2} \sqrt{2m\,[U(x) - \mathcal{E}_x]}\, dx = \int_0^{x_2} \sqrt{2m\,(C - eEx - \mathcal{E}_x)}\, dx =$$

$$= \frac{2}{3} \sqrt{2m} \frac{(C - \mathcal{E}_x)^{3/2}}{eE}.$$

Ainsi, le coefficient de pénétrabilité de la barrière pour des électrons d'énergie $\mathcal{E}_x$ est donné par la formule

$$D(\mathcal{E}_x) = D_0 e^{-\frac{4}{3}\frac{\sqrt{2m}}{\hbar}\frac{(C-\mathcal{E}_x)^{3/2}}{eE}}. \tag{29.2}$$

Suivant la valeur de $\mathcal{E}_x$ ce coefficient a des valeurs quelque peu différentes. On peut introduire un *coefficient de pénétrabilité moyen* ou *efficace* en prenant la moyenne sur $\mathcal{E}_x$ (afin que le courant d'émission soit le même). Tout moyennage se ramène au calcul de la moyenne des expressions de la forme $D_0 e^{-f(\mathcal{E}_x)/E}$, où la signification de la fonction $f(\mathcal{E}_x)$ apparaît clairement par comparaison avec la formule (29.2). Comme le calcul de la moyenne sur $\mathcal{E}_x$ se fait pour E fixé, il s'agit en fait du calcul de la moyenne de $f(\mathcal{E}_x)$. C'est une fonction positive puisque $C > \mathcal{E}$; par suite, après moyennage, on

peut la représenter sous forme d'une exponentielle. Finalement on trouve que le coefficient de pénétrabilité moyen est

$$\overline{D} = \overline{D}_0 e^{-E_0/E}, \qquad (29.3)$$

où $\overline{D}_0$ et E_0 sont des constantes dont les valeurs dépendent de la nature du métal. Le courant d'émission froide est donné par la formule

$$I(E) = I_0 \overline{D} = A e^{-E_0/E}. \qquad (29.4)$$

P. I. Loukirski confirma par l'expérience cette forme de la dépendance du courant d'émission froide avec le champ électrique extérieur.

Dans le problème ci-dessous on trouvera une comparaison entre la théorie classique de l'émission froide et les données expérimentales.

PROBLÈME

Une explication classique de l'émission d'électrons par une cathode froide faisait intervenir les forces d'image électrique $e^2/4x$ sous l'action desquelles les électrons seraient attirés vers la surface du métal (cf. t. III, § 23, 2). Compte tenu de cette force, l'énergie potentielle d'un électron se trouvant à une distance x de la surface du métal serait égale à $U = C - eEx - e^2/4x$. La prise en compte de cette force abaisse la barrière de potentiel que doit surmonter l'électron pour sortir du métal, i.e. réduit son travail de sortie. Calculer cette diminution du travail de sortie et évaluer l'intensité minimale du champ électrique extérieur à partir de laquelle pourrait se produire une émission froide. Faire le calcul numérique pour le tungstène, dont le travail de sortie $A = 4{,}5$ eV.

S o l u t i o n. La fonction U atteint son maximum pour $x = {}^1/_2 \sqrt{e/E}$, $U_{\max} = C - \sqrt{Ee^3}$. Ce résultat montre que la force d'image électrique diminue la hauteur de la barrière et donc le travail de sortie de $\sqrt{Ee^3}$. Le travail de sortie devient donc égal à $A' = A - \sqrt{Ee^3}$. L'émission froide débute lorsque $A' = 0$. Par conséquent,

$$E = \frac{A^2}{e^3} = \frac{V^2}{e}. \qquad (29.5)$$

On déduit (29.5) de l'expression précédente en exprimant le travail de sortie en fonction de la tension correspondante V à l'aide de la formule $A = eV$. Pour le tungstène, $V = 4{,}5$ V $= 1{,}5 \cdot 10^{-2}$ un. CGSE,

$$E = \frac{2{,}25 \cdot 10^{-4}}{4{,}8 \cdot 10^{-10}} = 0{,}469 \cdot 10^6 \text{ un. CGSE} = 1{,}41 \cdot 10^8 \text{ V/cm}.$$

Or Millikan obtenait de forts courants d'émission froide déjà pour $E \sim 4 \cdot 10^6$ V/cm.

CHAPITRE V

NOUVELLES ÉTAPES DE L'ÉLABORATION DE LA MÉCANIQUE QUANTIQUE ET SPECTRES

§ 30. La méthode des opérateurs

1. Comme on utilise largement la *méthode des opérateurs* en mécanique quantique, il convient d'en donner un bref aperçu. Cela est d'ailleurs indispensable pour concrétiser la forme des liens qui unissent le symbolisme de la mécanique quantique aux grandeurs observables.

On appelle *opérateur* un symbole mathématique qui, agissant sur une fonction de certaines variables, produit une autre fonction de ces mêmes variables. Un exemple d'une action de cette sorte est la multiplication par x ou par une fonction $f(x)$. En se plaçant à ce point de vue, on dira que les symboles x et $f(x)$ sont des opérateurs. Pour les distinguer des nombres on les dénote par $\hat{x}$ et $\hat{f}(x)$, i.e. on place un accent circonflexe au-dessus des symboles. Un autre exemple d'opérateur est la différentiation par rapport à x, i.e. $\frac{\partial}{\partial x}$, $\frac{\partial^2}{\partial x^2}$, . . .

On peut additionner les opérateurs. On entend par *somme d'opérateurs* $\hat{A}+\hat{B}$ un nouvel opérateur qui, agissant sur une fonction quelconque $f(x)$, fournit le résultat $\hat{A}f(x)+\hat{B}f(x)$. On entend par *produit d'opérateurs* $\hat{A}\hat{B}$ un nouvel opérateur dont l'action sur une fonction $f(x)$ se traduit par $\hat{A}\,[\hat{B}\,f(x)]$. Ici la fonction $f(x)$ est soumise d'abord à l'action de l'opérateur $\hat{B}$ et le résultat obtenu est ensuite soumis à l'action de l'opérateur $\hat{A}$. Un cas particulier de produit d'opérateurs est le produit d'un opérateur $\hat{A}$ par un nombre λ, i.e. soit $\lambda\hat{A}$, soit $\hat{A}\lambda$, puisqu'on peut considérer tout nombre comme un opérateur particulier. L'algèbre des opérateurs ne respecte pas toujours la loi de commutation en ce qui concerne la multiplication, autrement dit on n'a pas nécessairement $\hat{A}\hat{B}=\hat{B}\hat{A}$. Si cette égalité est vérifiée, on dit que les opérateurs $\hat{A}$ et $\hat{B}$ sont *commutatifs*. Dans le cas contraire, on dit que les opérateurs *ne sont pas commutatifs* ou encore *anticommutatifs*; un exemple en est la multiplication par x et la différentiation par rapport à x. On a en effet

$$\left(x\frac{\partial}{\partial x}\right)f=x\frac{\partial f}{\partial x},\quad \frac{\partial}{\partial x}\,(xf)=f+x\frac{\partial f}{\partial x},$$

de sorte que

$$\frac{\partial}{\partial x} x - x \frac{\partial}{\partial x} = 1. \tag{30.1}$$

A l'aide de ces définitions, on peut construire à partir d'opérateurs $\hat{A}$ et $\hat{B}$ donnés d'autres opérateurs $\hat{L}(\hat{A}, \hat{B})$ qui en sont des fonctions. Cette définition ne vaut que pour des *fonctions rationnelles entières* des opérateurs $\hat{A}$ et $\hat{B}$. Dans cette construction, la limitation imposée est suffisante pour la raison qu'en physique classique c'est dans les conditions d'une telle limitation qu'on arrive à définir de nouvelles grandeurs physiques à partir d'autres grandeurs physiques déjà définies.

L'addition et la multiplication des opérateurs se font d'après les règles algébriques usuelles d'addition et de multiplication des nombres. La seule différence réside en ce que dans une multiplication d'opérateurs on ne peut pas toujours permuter l'ordre des facteurs. Par exemple, on a toujours

$$(\hat{A} + \hat{B})^2 = (\hat{A} + \hat{B})(\hat{A} + \hat{B}) = \hat{A}^2 + \hat{B}\hat{A} + \hat{A}\hat{B} + \hat{B}^2,$$

et il serait incorrect d'écrire

$$(\hat{A} + \hat{B})^2 = \hat{A}^2 + 2\hat{A}\hat{B} + \hat{B}^2.$$

Cette dernière formule n'est juste que dans le cas où les opérateurs $\hat{A}$ et $\hat{B}$ commutent, puisque si $\hat{B}\hat{A} = \hat{A}\hat{B}$, elle se déduit de la formule précédente. Si les opérateurs ne sont pas commutatifs, $\hat{B}\hat{A} \neq \hat{A}\hat{B}$ et la formule ci-dessus est fausse.

L'opérateur $\hat{A}$ est dit *linéaire* si, agissant sur deux fonctions quelconques f et φ avec des constantes λ et μ arbitraires, il satisfait à la relation

$$\hat{A}(\lambda f + \mu\varphi) = \lambda\hat{A}f + \mu\hat{A}\varphi.$$

En mécanique quantique, on n'utilise que des opérateurs linéaires, car autrement le *principe de superposition des états* serait en défaut.

2. Supposons maintenant qu'on mesure la coordonnée x d'une particule un grand nombre de fois et qu'à chaque fois, dans la mesure où les conditions expérimentales le permettent, la particule est replacée dans les *mêmes conditions macroscopiques*. On peut alors caractériser l'état de la particule dans ces expériences par la fonction d'onde $\Psi(x)$ que nous supposerons, pour simplifier, ne dépendre que de la seule coordonnée spatiale x. La valeur moyenne de la coordonnée, qui sera déterminée à la suite de ces mesures, peut s'écrire sous la forme

$$\langle x \rangle = \int x \Psi^*\Psi \, dx,$$

puisque $\Psi^*\Psi\,dx$ est la probabilité de présence de la particule dans l'intervalle x, $x + dx$. On doit spécifier que la fonction $\Psi(x)$ est partout finie, différente de zéro dans une région limitée de l'espace et normalisée à l'unité, i.e.

$$\int \Psi^*\Psi\,dx = 1, \tag{30.2}$$

où l'intégration est étendue à tout l'espace, où Ψ est différent de zéro. Nous préférons écrire l'expression de la valeur moyenne $\langle x\rangle$ sous la forme

$$\langle x\rangle = \int \Psi^*\hat{x}\,\Psi\,dx. \tag{30.3}$$

La valeur moyenne de la fonction $f(x)$ se calcule exactement de la même manière, i.e. par la formule

$$\langle f(x)\rangle = \int \Psi^*(x)\,\hat{f}(x)\,\Psi(x)\,dx, \tag{30.4}$$

où $\hat{f}(x)$ est considéré comme un opérateur.

Dans le cas où l'état Ψ varie dans le temps, la formule (30.4) donne la valeur moyenne à un instant déterminé. Dans ce cas, dans la fonction $\Psi(x, t)\,t$ doit être considéré comme un *paramètre*, ce qui implique que t est une constante lors de l'intégration.

3. Voyons comment on fait pour calculer la valeur moyenne de l'impulsion d'une particule ou la valeur moyenne des fonctions rationnelles entières de l'impulsion, connaissant la fonction d'onde $\Psi(x)$. Admettons que la fonction $\Psi(x)$ est à tout instant partout finie et différente de zéro dans une région limitée de l'espace, ce qui implique qu'elle peut être normalisée par la formule (30.2). Pour simplifier mathématiquement le problème, on usera d'un artifice. Remplaçons la fonction d'onde $\Psi(x)$ vraie par une fonction périodique $\Phi(x)$ de période l, qui est telle que pour tout x $\Phi(x + l) = \Phi(x)$. La fonction $\Psi(x)$ n'est différente de zéro que sur un petit segment situé quelque part au milieu de l'intervalle $0 < x < l$ (donc de la période fondamentale). Dans cet intervalle, les deux fonctions $\Psi(x)$ et $\Phi(x)$ coïncident. En dehors de l'intervalle $0 < x < l$ la fonction $\Psi(x)$ s'annule. Par conséquent, la normalisation (30.2) entraîne la normalisation pour la fonction $\Phi(x)$:

$$\int_0^l \Phi^*(x)\,\Phi(x)\,dx = 1. \tag{30.5}$$

Ce changement de fonction ne peut affecter notablement les phénomènes qui évoluent à un instant donné dans une région donnée de l'espace. En effet, la différence qui existe entre les fonctions $\Phi(x)$ et $\Psi(x)$ ne se manifeste que dans des parties lointaines de l'espace qui ne peuvent exercer une influence notable sur les phénomènes étu-

diés. A la limite, lorsque $l \to \infty$, les petites distorsions des phénomènes résultant du remplacement de Ψ par Φ s'estompent entièrement.

Développons la fonction périodique $\Phi(x)$ en une série de Fourier (1768-1830):

$$\Phi(x) = \sum_{n=-\infty}^{n=+\infty} c_n e^{ik_n x}, \tag{30.6}$$

où

$$k_n = \frac{2\pi}{l} n. \tag{30.7}$$

Pour déterminer le coefficient c_m de cette série, il faut multiplier les deux membres de (30.6) par $e^{-ik_m x}$ et intégrer sur x de 0 à l. On trouve alors

$$\int_0^l e^{i(k_n - k_m)x}\, dx = \begin{cases} l \quad \text{lorsque} \quad n = m, \\ \dfrac{1}{i(k_n - k_m)} e^{i(k_n - k_m)x}\Big|_0^l = 0 \text{ pour } n \neq m. \end{cases}$$

Comme

$$e^{i(k_n - k_m)l} = e^{i2\pi(n-m)} = 1 = e^0,$$

on obtient

$$c_m = \frac{1}{l}\int_0^l \Phi(x)\, e^{-ik_m x}\, dx = \frac{1}{l}\int_0^l \Psi(x)\, e^{-ik_m x}\, dx. \tag{30.8}$$

Compte tenu de cette même condition, la normalisation se laisse ramener à la forme

$$\int_0^l \Phi^*\Phi\, dx = \int_0^l \Psi^*\Psi\, dx = l \sum_{n=-\infty}^{n=+\infty} |c_n|^2 = 1. \tag{30.9}$$

Etablissons encore une formule auxiliaire. On a

$$\int_0^l \Phi^* \frac{\partial}{\partial x} \Phi\, dx = \int_0^l \sum_m \sum_n c_m^* c_n e^{-ik_m x} \frac{\partial}{\partial x} e^{ik_n x}\, dx =$$

$$= i \int_0^l \sum_m \sum_n c_m^* c_n k_n e^{i(k_n - k_m)x}\, dx,$$

ou après permutation de l'ordre de la sommation et de l'intégration

$$\int_0^l \Phi^* \frac{\partial}{\partial x} \Phi\, dx = i \sum_m \sum_n c_m^* c_n k_n \int_0^l e^{i(k_n - k_m)x}\, dx.$$

Comme l'intégrale a déjà été calculée ci-dessus, on obtient

$$-i\int_0^l \Phi^* \frac{\partial}{\partial x}\Phi\, dx = l\sum |c_n|^2 k_n. \tag{30.10}$$

4. Jusqu'ici nous avons négligé la dépendance avec t de la fonction Ψ et par suite celle de la fonction Φ. Les calculs donnés ci-dessus concernaient les fonctions $\Psi(x, t)$ et $\Phi(x, t)$, où t était considéré comme un paramètre ayant une valeur fixe. On détermine la dépendance temporelle en exigeant que les fonctions Ψ et Φ satisfassent à l'équation temporelle de Schrödinger (21.5). A cette condition satisfait la fonction

$$\Phi(x, t) = \sum_n c_n e^{i(k_n x - \omega_n t)}, \tag{30.11}$$

où les fréquences ω_n correspondent à la loi de dispersion $\omega_n = \omega_n(k_n)$ conformément à la formule (19.6). Cette série est le développement de la fonction $\Phi(x, t)$ suivant les ondes planes de de Broglie *).

A l'onde de de Broglie $e^{i(k_n x - \omega_n t)}$ correspond l'impulsion $p_n = \hbar k_n$. L'impulsion a des valeurs discrètes, mais ce caractère est artificiel et résulte du remplacement de la fonction d'onde réelle $\Psi(x, t)$ par la fonction périodique auxiliaire $\Phi(x, t)$. *Les vraies valeurs de l'impulsion sont continues.* En effet, plus la période l que l'on prend est longue, plus courts sont les intervalles séparant les valeurs adjacentes du spectre discret des impulsions. A la limite où $l \to \infty$, cet intervalle tend vers zéro et l'impulsion devient une quantité variant de façon continue. La mesure de l'impulsion dans l'état $\Phi(x, t)$ donne une des valeurs p_n. En vertu de la condition de normalisation (30.9), la probabilité d'apparition de cette valeur est égale à $l\,|c_n|^2$. Il s'ensuit que la valeur moyenne de l'impulsion, obtenue à la suite des mesures, sera égale à

$$\langle p\rangle = \sum l\,|c_n|^2 p_n, \tag{30.12}$$

ou bien, en vertu de la relation (30.10),

$$\langle p\rangle = \int_0^l \Phi^* \left(-i\hbar\frac{\partial}{\partial x}\right)\Phi\, dx.$$

*) Une fonction arbitraire $\Phi(x, t)$ périodique en x peut être développée en une série de Fourier suivant x, avec des coefficients qui sont fonctions de t. Selon une opinion erronée, on obtient ainsi le développement suivant les ondes planes de de Broglie. L'erreur consiste en ce que la fonction périodique $\Phi(x, t)$ ne peut être une fonction arbitraire, il faut qu'elle vérifie l'équation de Schrödinger (21.5). Cela impose des contraintes aux coefficients de développement en tant que fonctions du temps t.

En passant à la limite $l \to \infty$, on obtient la formule

$$\langle p \rangle = \int \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi \, dx, \tag{30.13}$$

où toute indétermination liée à l'introduction de la fonction auxiliaire Φ est éliminée. L'intégration est étendue ici à tout l'espace infini, puisque les grandes valeurs de x ne contribuent pas à la valeur de l'intégrale.

A l'aide d'un raisonnement analogue, on obtient sans peine pour n arbitraire entier et positif la formule

$$\langle p^n \rangle = \int \Psi^* \left(-i\hbar \frac{\partial}{\partial x} \right)^n \Psi \, dx, \tag{30.14}$$

et pour une fonction entière et rationnelle de l'impulsion la formule

$$\langle F(p) \rangle = \int \Psi^* F(\hat{p}) \Psi \, dx, \tag{30.15}$$

où $\hat{p}$ désigne l'opérateur

$$\hat{p} \equiv \hat{p}_x = -i\hbar \frac{\partial}{\partial x}, \tag{30.16}$$

appelé *opérateur impulsion* ou plus exactement opérateur de la projection de l'impulsion $\hat{p}_x$.

5. La mécanique quantique postule que les résultats obtenus ci-dessus peuvent être étendus à toutes les grandeurs physiques qui sont des fonctions des coordonnées et des impulsions. Autrement dit, on pose que

$$\langle F(x, p) \rangle = \int \Psi^*(x) F(\hat{x}, \hat{p}) \Psi(x) \, dx, \tag{30.17}$$

où $F(x, p)$ est une fonction rationnelle entière des coordonnées et des impulsions, définie classiquement, et $F(\hat{x}, \hat{p})$ l'opérateur correspondant. La formule (30.17) constitue le fondement de l'introduction des opérateurs dans la mécanique quantique. Remarquons que puisque les opérateurs $\hat{x}$ et $\hat{p}$ sont rapportés à des coordonnées rectangulaires, pour trouver l'opérateur $F(\hat{x}, \hat{p})$ on doit utiliser la formule classique correspondante rapportée elle aussi aux coordonnées rectangulaires ou directement sous forme vectorielle.

Il convient de noter que x et p figurant dans la formule (30.17) ne peuvent désigner les valeurs de la coordonnée x et de l'impulsion p déterminées lors d'une même mesure. Une telle interprétation est contraire au principe d'incertitude de Heisenberg. On doit entendre par x et p la coordonnée et l'impulsion dans le sens classique. La mécanique quantique substitue à ces quantités les opérateurs $\hat{x}$ et $\hat{p}$ et introduit de nouveaux opérateurs $F(\hat{x}, \hat{p})$. On obtient l'opé-

rateur $F(\hat{x}, \hat{p})$ en appliquant aux opérateurs $\hat{x}$ et $\hat{p}$ les mêmes opérations de sommation et de multiplication que celles que l'on utilise en physique classique pour trouver les valeurs de la fonction $F(x, p)$ à partir des valeurs connues de x et p. Il n'y a encore aucun caractère statistique propre à la mécanique quantique. La statistique apparaît lorsqu'on passe à la formule (30.17) qui ne donne que la valeur moyenne de la fonction $F(x, p)$ et non sa vraie valeur (qui, en mécanique quantique, peut fort bien être dénuée de sens pour la raison qu'il est impossible de caractériser l'état d'une particule en fixant simultanément les valeurs de x et p). Le fait que l'opérateur $F(\hat{x}, \hat{p})$ se déduit de la fonction classique $F(x, p)$ établit, avec d'autres considérations, un lien étroit entre la mécanique classique et la mécanique quantique. On en arrive à une affirmation paradoxale, selon laquelle la justification de la mécanique quantique serait, *par principe, impossible* si on ne recourait pas à la mécanique classique, quoique la mécanique quantique soit une théorie plus générale, incorporant l'autre en tant que cas limite particulier. Ce cas limite se laisse déduire de la mécanique quantique chaque fois que la constante de Planck $\hbar$ est infiniment petite devant toutes les grandeurs de mêmes dimensions qui interviennent dans le phénomène considéré.

Tous les résultats ont été établis pour le cas unidimensionnel, dans le souci de simplifier et d'alléger les formules. Mais ces résultats se laissent aisément généraliser au cas tridimensionnel. C'est ainsi que l'opérateur de l'impulsion tridimensionnelle d'une particule est le vecteur symbolique

$$\hat{p} = -i\hbar\left(\frac{\partial}{\partial x}\,i + \frac{\partial}{\partial y}\,j + \frac{\partial}{\partial z}\,k\right) = -i\hbar\nabla, \qquad (30.18)$$

tandis que la formule de la valeur moyenne de cette impulsion est de la forme

$$\langle p\rangle = \int \Psi^* (-i\hbar\nabla)\,\Psi\,dV, \qquad (30.19)$$

où l'intégration est étendue à tout l'espace V et la fonction Ψ est supposée normalisée à l'unité:

$$\int \Psi^*\Psi\,dV = 1. \qquad (30.20)$$

Par conséquent, à toute quantité classique $F(\boldsymbol{r}, \boldsymbol{p})$ la mécanique quantique fait correspondre l'opérateur $F(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}})$ qui s'obtient en remplaçant les quantités classiques $\boldsymbol{r}$ et $\boldsymbol{p}$ par les opérateurs correspondants $\hat{\boldsymbol{r}}$ et $\hat{\boldsymbol{p}}$. Le lien avec les grandeurs réellement observables est établi sous une forme statistique à l'aide de la formule

$$\langle F(\boldsymbol{r}, \boldsymbol{p})\rangle = \int \Psi^* F(\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}})\,\Psi\,dV. \qquad (30.21)$$

6. Il importe de savoir s'il ne pourrait exister des états tels qu'en mesurant une quantité L correspondant à l'opérateur $\hat{L}$, on obtiendrait toujours une valeur L bien déterminée. On démontre facilement qu'à cette condition satisfont les fonctions d'onde qui sont des solutions de l'équation

$$\hat{L}\,\Psi = L\Psi. \tag{30.22}$$

Dans ce cas, on a en effet

$$\langle L\rangle = \int \Psi^*\hat{L}\Psi\, dx = \int \Psi^* L\Psi\, dx = L \int \Psi^*\Psi\, dx = L,$$

i.e. la valeur moyenne $\langle L\rangle$ est toujours égale à L. Or cela est possible si et seulement si chaque mesure fournit la valeur L. Nous avons démontré que la condition (30.22) était suffisante. Il n'est pas beaucoup plus difficile de démontrer qu'elle est aussi nécessaire, mais nous nous abstiendrons de le faire.

Les fonctions Ψ qui vérifient l'équation (30.22) sont appelées *fonctions propres* de l'opérateur $\hat{L}$, tandis que les nombres L en sont les *valeurs propres*. La mécanique quantique accepte que *toute mesure d'une grandeur physique ne peut fournir (avec une probabilité plus ou moins grande) que les valeurs propres de l'opérateur correspondant à cette grandeur*. Etant donné que les grandeurs physiques sont *essentiellement réelles*, leurs opérateurs $\hat{L}$ doivent être tels que les valeurs propres de ces opérateurs soient également *réelles*. On n'approfondira pas ici les conditions requises pour qu'il en soit ainsi.

L'équation (30.22) est une généralisation de la règle de quantification de l'énergie (cf. ch. IV) à n'importe quelles grandeurs physiques. Pour nous en assurer, cherchons l'opérateur $\hat{\mathscr{H}}$ correspondant à l'énergie totale d'une particule. D'après ce qui a été dit plus haut, cet opérateur doit être la somme de l'opérateur énergie cinétique et de l'opérateur énergie potentielle, i.e.

$$\hat{\mathscr{H}} = \frac{1}{2m}\hat{p}^2 + \hat{U} = -\frac{\hbar^2}{2m}\nabla^2 + U. \tag{30.23}$$

Par conséquent, (30.22) s'écrit dans ce cas sous la forme

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U\right)\psi = \mathscr{E}\psi.$$

Or c'est l'équation de Schrödinger (21.7) pour les états stationnaires. Sous forme symbolique, on peut l'écrire de façon concise

$$\hat{\mathscr{H}}\psi = \mathscr{E}\psi, \tag{30.24}$$

qui ne se distingue de (30.22) que par les notations utilisées. L'équation générale de Schrödinger (21.5) pour les états non stationnaires

peut aussi s'écrire sous forme symbolique :

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{\mathcal{H}} \Psi. \tag{30.25}$$

On notera que *l'équation* (30.25) *s'applique aussi bien au cas des forces dérivant d'un potentiel qu'à celui des forces qui n'en dérivent pas* (par exemple aux forces magnétiques). Il faut seulement que les équations classiques correspondantes puissent être récrites sous la forme des *équations de Hamilton* (1805-1865). Dans ce cas, l'opérateur $\hat{\mathcal{H}}$ est appelé *opérateur de Hamilton* ou *hamiltonien*. Si les forces dérivent d'un potentiel, l'hamiltonien devient identique à l'opérateur énergie.

Donnons un autre exemple d'application de l'équation (30.22). Il s'agit de trouver les fonctions propres et les valeurs propres de l'opérateur impulsion. Dans le cas unidimensionnel, on pose $\hat{L} = \hat{p} = -i\hbar\partial/\partial x$ et on obtient

$$-i\hbar \frac{\partial \Psi}{\partial x} = p\Psi,$$

d'où

$$\Psi = C(t)\,e^{\frac{ipx}{\hbar}} = C(t)\,e^{ikx}.$$

Pour que l'équation générale de Schrödinger (30.25) soit vérifiée, il convient de poser $C(t) = Ce^{-i\omega t}$, i.e.

$$\Psi = Ce^{i(kx-\omega t)}. \tag{30.26}$$

Ce résultat montre que les fonctions propres de l'opérateur impulsion sont des ondes planes de de Broglie. Le paramètre p peut prendre n'importe quelle valeur, ce qui signifie que le spectre des valeurs propres de l'opérateur $\hat{p}$ est continu.

Remarquons à ce propos que l'équation (30.22) a été déduite de l'équation (30.21) en admettant que la condition de normalisation (30.20) était satisfaite, i.e. la fonction Ψ s'annulait à l'infini. Les fonctions propres (30.26) ne satisfont pas à cette condition. Dans le cas de l'opérateur énergie, lorsque $\mathcal{E} > 0$, les valeurs propres constituent un spectre continu et la condition de normalisation (30.20) n'est pas satisfaite. Dans ces cas, la démonstration de la formule (30.22) donnée ci-dessus ne convient pas, quoique la formule (30.22) reste valable. On peut cependant généraliser la normalisation (30.20) afin que la démonstration de (30.22) soit valable dans tous les cas. Pour ce faire, il faut faire appel à des *fonctions généralisées*, mais on n'examinera pas ici cette question de caractère formel. En physique on peut se contenter, en principe, de fonctions d'onde s'annulant à l'infini, auxquelles la normalisation (30.20) s'applique toujours.

7. Pour conclure, examinons une question spécifique pour la mécanique quantique et qui ne se pose donc pas en mécanique classique. Soient $\hat{A}$ et $\hat{B}$ deux opérateurs quantiques ; à chacun d'eux correspond son spectre de valeurs propres. La question que l'on doit se poser est la suivante : existe-t-il toujours un état Ψ pour lequel les deux opérateurs ont des valeurs propres A et B bien déterminées? Autrement dit, on voudrait savoir s'il peut exister un état Ψ dans lequel les deux quantités A et B sont simultanément mesurables. Supposons que Ψ_n est la fonction propre des opérateurs $\hat{A}$ et $\hat{B}$, i.e.

$$\hat{A}\Psi_n = A_n\Psi_n, \qquad \hat{B}\Psi_n = B_n\Psi_n,$$

où A_n et B_n sont des nombres exprimant les valeurs propres des opérateurs $\hat{A}$ et $\hat{B}$ dans un seul et même état Ψ_n. Multiplions par l'opérateur $\hat{B}$ la première égalité de gauche :

$$\hat{B}\hat{A}\Psi_n = \hat{B}A_n\Psi_n = A_n\hat{B}\Psi_n = A_nB_n\Psi_n.$$

On trouve de façon analogue que

$$\hat{A}\hat{B}\Psi_n = B_nA_n\Psi_n.$$

De là $(\hat{A}\hat{B} - \hat{B}\hat{A})\,\Psi_n = 0$. Ce résultat ne nous permet pas de conclure que $\hat{A}\hat{B} - \hat{B}\hat{A} = 0$, puisque Ψ_n n'est pas une fonction quelconque, mais une des fonctions propres communes des opérateurs $\hat{A}$ et $\hat{B}$ *).

Supposons cependant que chaque fonction propre de l'opérateur $\hat{A}$ est une fonction propre de l'opérateur $\hat{B}$ et réciproquement. Il existe un théorème mathématique qu'on s'abstiendra de démontrer, qui affirme que toute fonction d'onde Ψ peut être développée suivant les fonctions propres de l'opérateur $\hat{A}$ (ou suivant celles de $\hat{B}$), i.e.

$$\Psi = \sum c_n\Psi_n$$

(cela suppose que le spectre est discret, ce qui n'a pas d'importance). Il résulte de cette formule et de la relation $(\hat{A}\hat{B} - \hat{B}\hat{A})\,\Psi_n = 0$ que

$$(\hat{A}\hat{B} - \hat{B}\hat{A})\,\Psi = 0.$$

Puisque Ψ est une fonction arbitraire, on est maintenant en droit de conclure que

$$\hat{A}\hat{B} = \hat{B}\hat{A}, \tag{30.27}$$

*) Par exemple, de l'égalité $\left(\frac{\partial^2}{\partial x^2} - \frac{\partial^3}{\partial x^3}\right)x = 0$ il ne s'ensuit pas que $\frac{\partial^2}{\partial x^2} - \frac{\partial^3}{\partial x^3} = 0$.

ce qui signifie que les opérateurs $\hat{A}$ et $\hat{B}$ sont commutatifs. En effet, le seul opérateur susceptible de rendre nulle n'importe quelle fonction est l'opérateur multiplication par zéro.

On peut donc énoncer : *si toutes les fonctions propres des opérateurs $\hat{A}$ et $\hat{B}$ coïncident, les opérateurs sont commutatifs.* Le théorème inverse est également vrai. *Si les opérateurs $\hat{A}$ et $\hat{B}$ sont commutatifs, leurs fonctions propres sont les mêmes.* On ne démontrera pas ce théorème.

Le théorème ci-dessus peut être formulé d'une autre façon. *Deux quantités A et B sont en général mesurables simultanément si et seulement si les opérateurs correspondants $\hat{A}$ et $\hat{B}$ sont commutatifs.* Cette règle n'est en défaut que dans de rares cas particuliers (cf. § 31, 1 et 4).

Par exemple, on peut mesurer simultanément les coordonnées x et y, car les opérateurs $\hat{x}$ et $\hat{y}$ sont commutatifs. Par contre, on ne peut mesurer simultanément la coordonnée x et l'impulsion correspondante p_x parce que les opérateurs $\hat{x}$ et $\hat{p}_x$ ne sont pas commutatifs, comme le montre la formule (30.1). C'est d'ailleurs la conclusion qu'on tire du principe d'incertitude de Heisenberg. Mais la coordonnée x et l'impulsion p_y correspondant à une autre coordonnée y sont simultanément mesurables puisque les opérateurs $\hat{x}$ et $\hat{p}_y = -i\hbar\partial/\partial y$ commutent entre eux pour la raison que la coordonnée x se comporte comme une constante lors de la dérivation par rapport à y.

§ 31. Moment cinétique d'une particule

1. En mécanique classique, le moment cinétique $\boldsymbol{l}$ d'une particule par rapport à l'origine des coordonnées O est défini par le produit vectoriel $[\boldsymbol{rp}]$. En mécanique quantique, cette définition n'a pas de sens puisqu'il n'existe pas d'état dans lequel les deux vecteurs $\boldsymbol{r}$ et $\boldsymbol{p}$ auraient des valeurs déterminées. En mécanique quantique on fait correspondre au produit vectoriel $[\boldsymbol{rp}]$ l'opérateur

$$\hat{\boldsymbol{l}} = [\hat{\boldsymbol{r}}\hat{\boldsymbol{p}}]. \tag{31.1}$$

En développant ce produit vectoriel et en respectant l'ordre de disposition des opérateurs des coordonnées et des projections du vecteur impulsion, on trouve les opérateurs des projections du moment cinétique sur les axes de coordonnées X, Y, Z :

$$\begin{aligned}
\hat{l}_x &= (\hat{y}\hat{p}_z - \hat{z}\hat{p}_y) = i\hbar\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right),\\
\hat{l}_y &= (\hat{z}\hat{p}_x - \hat{x}\hat{p}_z) = i\hbar\left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right),\\
\hat{l}_z &= (\hat{x}\hat{p}_y - \hat{y}\hat{p}_x) = i\hbar\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right).
\end{aligned} \tag{31.2}$$

En fonction de ces projections, l'opérateur moment cinétique se laisse exprimer par la formule

$$\hat{l} = i\hat{l}_x + j\hat{l}_y + k\hat{l}_z. \tag{31.3}$$

Afin de dégager la signification de cet opérateur vectoriel, il suffit d'écrire le résultat de son action sur une fonction arbitraire ψ. On trouve alors

$$\hat{l}\psi = i(\hat{l}_x\psi) + j(\hat{l}_y\psi) + k(\hat{l}_z\psi), \tag{31.4}$$

donc un vecteur de composantes $\hat{l}_x\psi$, $\hat{l}_y\psi$ et $\hat{l}_z\psi$. Ainsi, à une fonction d'onde ψ quelconque correspond un vecteur défini par l'expression (31.4). On peut cependant se demander s'il existe une fonction ψ pour laquelle les trois projections du vecteur (31.4) auraient des valeurs déterminées, c.-à-d. qu'on ait simultanément les trois égalités

$$\hat{l}_x\psi = l_x\psi, \quad \hat{l}_y\psi = l_y\psi, \quad \hat{l}_z\psi = l_z\psi. \tag{31.5}$$

Pour donner une réponse à cette question, il faut trouver les règles de commutation des opérateurs $\hat{l}_x$, $\hat{l}_y$, $\hat{l}_z$. En multipliant l'un par l'autre les opérateurs $\hat{l}_x$ et $\hat{l}_y$ et en respectant l'ordre de leur disposition, on obtient

$$\hat{l}_x\hat{l}_y = -\hbar^2\left(z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}\right)\left(x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}\right) =$$

$$= -\hbar^2\left(z\frac{\partial}{\partial y}x\frac{\partial}{\partial z} - y\frac{\partial}{\partial z}x\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}z\frac{\partial}{\partial x} + y\frac{\partial}{\partial z}z\frac{\partial}{\partial x}\right) =$$

$$= -\hbar^2\left(zx\frac{\partial^2}{\partial y\,\partial z} - yx\frac{\partial^2}{\partial z^2} - z^2\frac{\partial^2}{\partial y\,\partial x} + yz\frac{\partial^2}{\partial z\,\partial x} + y\frac{\partial}{\partial x}\right).$$

On trouve de même

$$\hat{l}_y\hat{l}_x = -\hbar^2\left(zy\frac{\partial^2}{\partial x\,\partial z} - xy\frac{\partial^2}{\partial z^2} - z^2\frac{\partial^2}{\partial x\,\partial y} + xz\frac{\partial^2}{\partial z\,\partial y} + x\frac{\partial}{\partial y}\right).$$

Les opérations de différentiation par rapport à deux variables indépendantes sont commutatives, i.e. par exemple $\frac{\partial^2}{\partial x\,\partial y} = \frac{\partial^2}{\partial y\,\partial x}$. Il s'ensuit qu'en soustrayant membre à membre les deux égalités ci-dessus, on trouvera

$$\hat{l}_x\hat{l}_y - \hat{l}_y\hat{l}_x = -\hbar^2\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) = i\hbar\hat{l}_z.$$

En procédant de la même façon, on trouve les deux autres règles de commutation. On obtient ainsi

$$\begin{aligned} \hat{l}_y\hat{l}_z - \hat{l}_z\hat{l}_y &= i\hbar\hat{l}_x, \\ \hat{l}_z\hat{l}_x - \hat{l}_x\hat{l}_z &= i\hbar\hat{l}_y, \\ \hat{l}_x\hat{l}_y - \hat{l}_y\hat{l}_x &= i\hbar\hat{l}_z. \end{aligned} \tag{31.6}$$

Ainsi, *deux projections quelconques de l'opérateur moment ne sont pas commutatives. Par suite, il n'existe aucun état dans lequel les trois projections ou deux seulement des trois projections* l_x, l_y, l_z *auraient des valeurs déterminées* (cf. § 30, 7). La seule exception à cette règle est le cas où les trois projections sont simultanément nulles. *Il ne peut donc exister d'état dans lequel le vecteur moment cinétique lui-même*

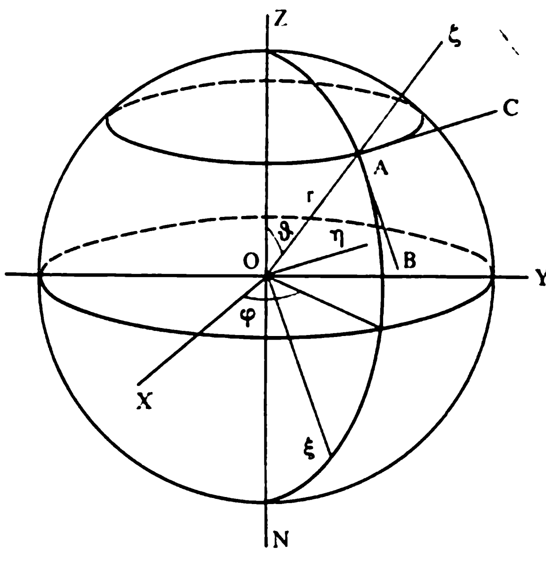


Fig. 57

aurait une valeur déterminée, i.e. serait complètement défini en grandeur et en direction. Autrement dit, *l'opérateur moment* $\hat{l}$ *ne possède ni fonctions propres ni valeurs propres vectorielles correspondantes.* Le lien entre l'opérateur vectoriel moment et la réalité physique se laisse établir sous une forme générale *statistiquement* — à l'aide de la formule (30.21) servant à déterminer les valeurs moyennes obtenues par mesure du moment cinétique lui-même et non pas de l'opérateur correspondant.

2. Par quelles grandeurs physiques (et non par leurs opérateurs) caractérise-t-on en mécanique quantique le moment cinétique d'une particule? Pour trouver une réponse à cette question, exprimons d'abord les opérateurs des projections du moment cinétique en coordonnées polaires. On peut le faire simplement par la méthode analytique, mais ce serait assez laborieux. Il est préférable d'utiliser la méthode géométrique. Plaçons au point commun O l'origine des coordonnées du système rectangulaire XYZ et du système polaire (fig. 57). En coordonnées polaires, la position d'un point A est ca-

ractérisée par sa distance r au point O et par deux angles: ϑ (angle polaire) et φ (azimut). Menons par le point A la tangente AB au méridien ZAN et la tangente AC au parallèle correspondant. Introduisons un système de coordonnées rectangulaires auxiliaires ξ, η, ζ dont les axes $O\xi$ et $O\eta$ sont respectivement parallèles à AB et AC, l'axe $O\zeta$ étant dirigé suivant le rayon OA. En supposant que la particule (classique) se trouve au point A, ses coordonnées seront $\xi = \eta = 0$, $\zeta = r$. En remplaçant x, y, z par ξ, η, ζ dans les formules (31.2), on écrira

$$\hat{l}_\xi = i\hbar r \frac{\partial}{\partial\eta}, \quad \hat{l}_\eta = -i\hbar r \frac{\partial}{\partial\xi}, \quad \hat{l}_\zeta = 0.$$

Or, d'après la figure 57, $d\xi = r d\vartheta$ (pour r et η constants), $d\eta = r \sin\vartheta \, d\varphi$ (pour r et ξ constants). Par conséquent,

$$\hat{l}_\xi = i \frac{\hbar}{\sin\vartheta} \frac{\partial}{\partial\varphi}, \quad \hat{l}_\eta = -i\hbar \frac{\partial}{\partial\vartheta}, \quad \hat{l}_\zeta = 0.$$

Maintenant il est facile de passer au système initial de coordonnées rectangulaires XYZ en utilisant la règle usuelle de transformation des projections d'un vecteur. La possibilité d'appliquer cette règle aux opérateurs est due à ce que les cosinus directeurs définissant les positions relatives des systèmes de coordonnées XYZ et $\xi\eta\zeta$ doivent être considérés comme des quantités constantes sur lesquelles les opérateurs $\frac{\partial}{\partial\vartheta}$ et $\frac{\partial}{\partial\varphi}$ n'exercent aucune action. Ecrivons ces cosinus directeurs dans le tableau suivant:

	ξ	η	ζ
X	$\cos\vartheta\cos\varphi$	$-\sin\varphi$	
Y	$\cos\vartheta\sin\varphi$	$\cos\varphi$	
Z	$-\sin\vartheta$		$\cos\vartheta$

On a alors

$$\hat{l}_x = \hat{l}_\xi \cos\vartheta\cos\varphi - \hat{l}_\eta \sin\varphi,$$

$$\hat{l}_y = \hat{l}_\xi \cos\vartheta\sin\varphi + \hat{l}_\eta \cos\varphi,$$

$$\hat{l}_z = -\hat{l}_\xi \sin\vartheta$$

ou

$$\hat{l}_x = i\hbar \left(\sin\varphi \frac{\partial}{\partial\vartheta} + \operatorname{cotg}\vartheta \cos\varphi \frac{\partial}{\partial\varphi}\right),$$
$$\hat{l}_y = i\hbar \left(-\cos\varphi \frac{\partial}{\partial\vartheta} + \operatorname{cotg}\vartheta \sin\varphi \frac{\partial}{\partial\varphi}\right), \qquad (31.7)$$
$$\hat{l}_z = -i\hbar \frac{\partial}{\partial\varphi}.$$

Il convient de noter la différence qui existe entre le moment cinétique classique $[\boldsymbol{rp}]$ d'une particule et l'opérateur moment de la mécanique quantique défini par (31.3). Le moment classique dépend du rayon vecteur $\boldsymbol{r}$, donc du choix de l'origine des coordonnées O par rapport à laquelle on le détermine. La formule (31.7) montre que l'opérateur moment ne dépend pas de $\boldsymbol{r}$, mais dépend des angles ϑ et φ. Cela signifie que l'opérateur moment (31.3) ne dépend pas du choix de l'origine des coordonnées et dépend uniquement de l'orientation des axes de coordonnées. Il serait donc plus exact de l'appeler *opérateur moment angulaire* ou *opérateur moment rotatoire* de la particule. A la différence du moment cinétique classique, il est inutile d'indiquer par rapport à quelle origine il est défini. Bien entendu, les valeurs propres des opérateurs des projections et du carré du moment angulaire (il en sera question ci-après) sont également indépendantes du choix de l'origine des coordonnées.

3. Une question s'impose aussitôt à l'esprit : une des projections du moment angulaire peut-elle avoir une valeur déterminée ? On conçoit que si la réponse à cette question est positive (ou négative) pour l'une des projections, elle l'est aussi pour les deux autres, ainsi que pour la projection du moment angulaire sur un axe de direction arbitraire. Cette assertion résulte directement de l'isotropie de l'espace, i.e. de l'équivalence de toutes les directions spatiales. Pour un même état, les projections du moment angulaire sur deux axes différents ne peuvent avoir des valeurs déterminées. Il s'ensuit que le choix de la direction est arbitraire. En général, c'est la direction de l'axe Z, puisque l'expression de l'opérateur $\hat{l}_z$ est la plus simple de toutes. La réponse à la question posée se fonde sur l'équation

$$\hat{l}_z\psi = l_z\psi$$

ou

$$-i\hbar \frac{\partial\psi}{\partial\varphi} = l_z\psi, \qquad (31.8)$$

où l_z est une constante. Cette équation a pour solution

$$\psi = C(r, \vartheta)\, e^{i\frac{l_z}{\hbar}\varphi}. \qquad (31.9)$$

Puisqu'on impose que ψ soit univoque, cette fonction doit reprendre sa valeur initiale lorsque l'argument φ s'accroît de 2π puisque alors le point considéré retrouve sa position initiale. On doit donc avoir

$$e^{i\frac{l_z}{\hbar}\varphi} = e^{i\frac{l_z}{\hbar}(\varphi+2\pi)}.$$

Comme la fonction exponentielle est périodique de période $2\pi i$, cette égalité n'est vérifiée que si

$$i\frac{l_z}{\hbar}\cdot 2\pi = m\cdot 2\pi i$$

ou

$$l_z = m\hbar, \tag{31.10}$$

où m est un nombre entier positif ou négatif, y compris le zéro ($m = 0, \pm 1, \pm 2, \ldots$). L'égalité (31.10) exprime le fait que la projection du moment angulaire sur un axe de direction arbitraire est nécessairement quantifiée. Pour simplifier encore cette formule, on prend pour unité de moment angulaire la constante $\hbar$. Avec cette convention, la projection du moment angulaire sur une direction donnée ne peut prendre que des valeurs entières $m = 0, \pm 1, \pm 2, \ldots$ On note m_z la projection du moment angulaire sur l'axe Z, par conséquent $m_z = m$.

La règle (31.10) se présente sous une forme analogue à celle de la règle de quantification correspondante (13.6) du moment cinétique de la théorie de Bohr. Ces deux règles sont cependant fort différentes. Dans (13.6) on doit entendre par L le moment cinétique total de la particule (de l'électron), tandis que dans (31.10) il ne s'agit que d'une projection du moment cinétique sur une direction quelconque, le vecteur moment cinétique en tant que quantité bien définie étant inexistant.

4. On peut s'assurer directement que la fonction propre de l'opérateur $\hat{l}_z$, i.e. la fonction (31.9), ne peut être simultanément la fonction propre de l'opérateur $\hat{l}_x$ ou de l'opérateur $\hat{l}_y$. Supposons qu'il n'en soit pas ainsi et que la fonction (31.9) est la fonction propre de l'opérateur $\hat{l}_x$ par exemple. On aurait alors

$$\hat{l}_x\psi = i\hbar\left(\sin\varphi\frac{\partial C}{\partial\vartheta} + im\operatorname{cotg}\vartheta\cos\varphi C\right)e^{im\varphi}.$$

Ce résultat ne peut être réduit à la forme $l_x\psi$ (l_x étant un nombre constant). Pour s'en assurer, il suffit d'envisager le cas particulier où $\varphi = 0$. On a alors

$$\hat{l}_x\psi = -\hbar m\operatorname{cotg}\vartheta\psi.$$

D'autre part, d'après l'hypothèse faite, on devrait avoir

$$\hat{l}_x\psi = l_x\psi.$$

Ces deux relations devraient être vérifiées pour toute valeur de ϑ, ce qui est évidemment impossible. Le même raisonnement appliqué à l'opérateur $\hat{l}_y$ conduit à la même conclusion. Ainsi se trouve confirmée l'assertion qu'il ne peut exister de fonction d'onde Ψ qui soit simultanément fonction propre des opérateurs $\hat{l}_x$, $\hat{l}_y$, $\hat{l}_z$. La seule exception concerne le cas où la fonction Ψ présente une symétrie sphérique, i.e. ne dépend que de r. Dans ce dernier cas, Ψ sera fonction propre des trois opérateurs $\hat{l}_x$, $\hat{l}_y$, $\hat{l}_z$ et les trois projections m_x, m_y, m_z du moment angulaire seront nulles.

5. La seconde quantité qui caractérise, à côté de la projection m_z, la valeur du moment angulaire est le *carré du moment angulaire total*. On le dénote généralement par l^2, quoique ce ne soit pas le carré du vecteur l (puisqu'il n'existe pas), mais la *valeur propre du carré de l'opérateur moment angulaire*, i.e.

$$l^2 = (\hat{l}_x i + \hat{l}_y j + \hat{l}_z k)^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2. \qquad (31.11)$$

On peut donc dire que la notation l^2 généralement en usage est décevante, mais on la conservera tout en recommandant au lecteur d'avoir toujours présent à l'esprit que cette notation désigne le carré de l'opérateur moment angulaire $\hat{l}^2$.

Afin de s'assurer que les quantités l^2 et m_z peuvent être mesurées pour un même état, il faut démontrer que les opérateurs $\hat{l}^2$ et $\hat{l}_z$ sont commutatifs. Ecrivons à cet effet

$$\hat{l}^2\hat{l}_z = (\hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2)\,\hat{l}_z = \hat{l}_x\,(\hat{l}_x\hat{l}_z) + \hat{l}_y\,(\hat{l}_y\hat{l}_z) + \hat{l}_z^3,$$

et en vertu des relations de commutation (31.6)

$$\hat{l}^2\hat{l}_z = \hat{l}_x\,(\hat{l}_z\hat{l}_x - i\hbar\hat{l}_y) + \hat{l}_y\,(\hat{l}_z\hat{l}_y + i\hbar\hat{l}_x) + \hat{l}_z^3.$$

De façon analogue

$$\hat{l}_z\hat{l}^2 = (\hat{l}_x\hat{l}_z + i\hbar\hat{l}_y)\,\hat{l}_x + (\hat{l}_y\hat{l}_z - i\hbar\hat{l}_x)\,\hat{l}_y + \hat{l}_z^3.$$

En soustrayant membre à membre, on trouve

$$\hat{l}^2\hat{l}_z - \hat{l}_z\hat{l}^2 = 0, \qquad (31.12)$$

c'est ce qu'il fallait démontrer. Il va de soi que la même relation de commutation est valable pour les opérateurs $\hat{l}_x$ et $\hat{l}_y$.

Il existe donc un état dans lequel le carré du moment angulaire l^2 et l'une de ses projections sur une direction choisie ont simultanément

des valeurs déterminées. La direction que l'on choisit généralement est l'axe Z.

Considérons un état quelconque Ψ dans lequel les quantités l^2 et m_z (et par suite m_z^2) ont simultanément des valeurs déterminées. Démontrons que dans cet état on a toujours $l^2 > m_z^2$ (excepté le cas où $l^2 = 0$). Examinons à cet effet l'égalité opératorielle

$$\hat{l}^2 - \hat{l}_z^2 = \hat{l}_x^2 + \hat{l}_y^2.$$

Pour l'opérateur $\hat{l}^2 - \hat{l}_z^2$ la fonction Ψ est la fonction propre ayant pour valeur propre $l^2 - m_z^2$. La valeur moyenne correspondante est égale à cette même valeur. En calculant la moyenne par la formule (30.21), on obtient

$$l^2 - m_z^2 = \int \Psi^* (\hat{l}_x^2 + \hat{l}_y^2) \Psi \, dV.$$

Comme l'intégrale du second membre est essentiellement positive, vu qu'elle représente la valeur moyenne de la quantité essentiellement positive $l_x^2 + l_y^2$ dans l'état Ψ, on a

$$l^2 > m_z^2. \tag{31.13}$$

Il s'ensuit que le moment angulaire ne peut s'orienter exactement le long de l'axe Z. Quel que soit l'état, il comporte toujours les projections l_x et l_y qui, dans l'état considéré, sont toujours indéterminées. C'est là un résultat qu'on connaît déjà : il n'existe pas d'état dans lequel les trois projections l_x, l_y, l_z puissent avoir des valeurs déterminées.

On dit parfois que dans l'état propre où l^2 a une valeur déterminée, le moment angulaire conserve sa longueur égale à $\sqrt{l^2}$, mais que sa direction subit des fluctuations. Il nous semble que cet énoncé n'est pas très heureux, puisqu'il suppose l'existence d'un vecteur l ayant à tout instant une longueur et une direction déterminées, mais cette direction change de façon désordonnée et continue au cours du temps. Or, il a été prouvé ci-dessus qu'un tel vecteur n'existe pas et que, par conséquent, il ne peut être question de ses fluctuations.

6. Il nous reste à déterminer les valeurs propres l^2 de l'opérateur du carré du moment angulaire $\hat{l}^2$. Pour cela il suffit d'appliquer les règles de commutation (31.6). Mais auparavant il convient d'exprimer ces règles sous une forme plus commode. Introduisons deux opérateurs

$$\hat{l}_+ = \hat{l}_x + i\hat{l}_y, \quad \hat{l}_- = \hat{l}_x - i\hat{l}_y. \tag{31.14}$$

A l'aide de (31.6) il est aisé de trouver

$$\begin{aligned} \hat{l}_+\hat{l}_- - \hat{l}_-\hat{l}_+ &= 2\hbar\hat{l}_z, \\ \hat{l}_z\hat{l}_+ - \hat{l}_+\hat{l}_z &= \hbar\hat{l}_+, \\ \hat{l}_z\hat{l}_- - \hat{l}_-\hat{l}_z &= -\hbar\hat{l}_-. \end{aligned} \tag{31.15}$$

D'autre part,

$$\hat{l}^2 = \left(\frac{\hat{l}_+ + \hat{l}_-}{2}\right)^2 + \left(\frac{\hat{l}_+ - \hat{l}_-}{2i}\right)^2 + \hat{l}_z^2,$$

ou bien

$$\hat{l}^2 = \hat{l}_+\hat{l}_- + \hat{l}_z^2 - \hbar\hat{l}_z = \hat{l}_-\hat{l}_+ + \hat{l}_z^2 + \hbar\hat{l}_z. \tag{31.16}$$

Comme l^2 est une quantité bornée, il résulte de (31.13) que m_z^2 et donc m_z sont également des quantités bornées. Notons l la plus grande valeur positive de la projection m_z correspondant à une valeur donnée du carré du moment l^2. Soit Ψ la fonction d'onde commune aux opérateurs $\hat{l}^2$ et $\hat{l}_z$, la valeur propre de l'opérateur $\hat{l}_z$ étant égale à m. Sous cette hypothèse

$$\hat{l}^2\Psi = l^2\Psi, \quad \hat{l}_z\Psi = \hbar l\Psi. \tag{31.17}$$

Appliquons à cette fonction les relations de commutation (31.15):

$$\hat{l}_z\hat{l}_\pm\Psi = (\hat{l}_\pm\hat{l}_z \pm \hbar\hat{l}_\pm)\Psi = \hbar(l \pm 1)\, l_\pm\Psi.$$

Cette expression prouve que les fonctions $\hat{l}_+\Psi$ et $\hat{l}_-\Psi$ sont les fonctions propres de l'opérateur $\hat{l}_z$ ayant pour valeurs propres respectivement $\hbar(l+1)$ et $\hbar(l-1)$. Or, la quantité $\hbar(l+1)$ ne peut être une valeur propre de l'opérateur $\hat{l}_z$ puisque, par hypothèse, la plus grande valeur propre de cet opérateur est $\hbar l$. Par conséquent, l'égalité

$$\hat{l}_z\hat{l}_+\Psi = \hbar(l+1)\,\hat{l}_+\Psi$$

est impossible. Or, cette égalité se déduit logiquement des relations de commutation (31.15) et de l'équation $\hat{l}_z\Psi = l\Psi$. On ne peut lever cette contradiction que si $\hat{l}_+\Psi = 0$, puisque dans ce cas cette égalité serait satisfaite. Or, cela implique que $\hat{l}_-\hat{l}_+\Psi = 0$, ou en vertu de (31.16)

$$(\hat{l}^2 - \hat{l}_z^2 - \hbar\hat{l}_z)\Psi = 0.$$

Or, en vertu de (31.17) $\hat{l}_z^2\Psi = \hbar^2 l^2\Psi$, $\hbar\hat{l}_z\Psi = \hbar^2 l\Psi$, de sorte que

$$(\hat{l}^2 - \hbar^2 l^2 - \hbar^2 l)\Psi = 0,$$

ou

$$\hat{l}^2\Psi = \hbar^2 l(l+1)\Psi.$$

Il en résulte que Ψ est la fonction propre de l'opérateur du carré du moment angulaire, ayant pour valeur propre

$$l^2 = \hbar^2 l(l+1). \tag{31.18}$$

7. Posons que le carré du moment angulaire l^2 a une valeur déterminée $l(l+1)$. Pour combien d'états cela peut-il être vrai,

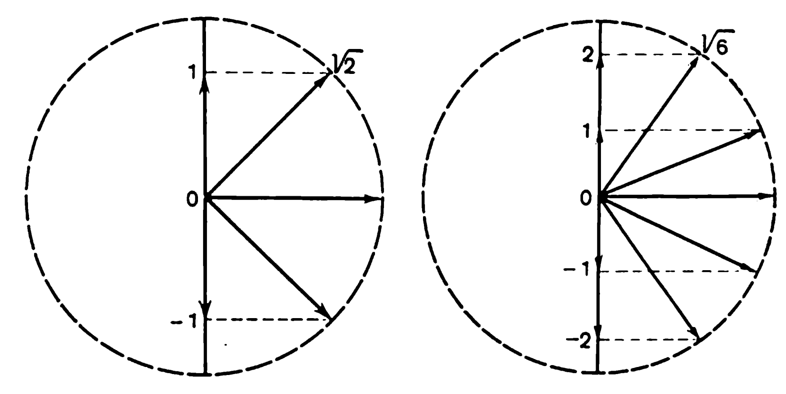


Fig. 58

sachant que dans ces états la projection m possède aussi une valeur déterminée? De toute évidence, dans ces états m peut prendre les valeurs suivantes:

$$m = -l, -(l-1), \ldots, -1, 0, +1, \ldots, +(l-1), +l,$$

ce qui singifie qu'il peut y avoir $(2l+1)$ états satisfaisant aux conditions imposées.

Ces résultats qui déterminent les valeurs possibles de l_z et l^2 sont désignés sous le nom de *quantification spatiale*. Ce terme a été emprunté à l'ancienne théorie de Bohr, où la quantification spatiale caractérisait les orientations spatiales possibles du vecteur moment angulaire l. Du point de vue de la mécanique quantique, le terme « quantification spatiale » n'est pas adéquat, puisque le soi-disant vecteur l ne peut avoir, par principe, d'orientations spatiales déterminées. Aux fins de visualisation, on représente généralement la quantification spatiale sous forme graphique à l'aide de *diagrammes vectoriels*. On porte sur l'axe Z les valeurs possibles de m en les considérant comme les projections du vecteur l de longueur $\sqrt{l(l+1)}$, ayant des orientations spatiales discrètes. A titre d'exemple, on a représenté sur la figure 58 les diagrammes vectoriels pour $l = 1$ et

$l = 2$ (l'unité de moment angulaire a été posée égale à la constante de Planck $\hbar$). On ne doit pas prendre ces diagrammes à la lettre. Ils ne reflètent correctement que deux faits réels: les valeurs possibles de la projection m et les valeurs possibles du carré du moment angulaire l^2. Pour des raisons qui seront précisées ultérieurement, le nombre quantique l est appelé *nombre quantique orbital* et le nombre m est appelé *nombre quantique magnétique*.

8. En mécanique classique, l'énergie cinétique d'un corps solide en rotation est donnée par la formule

$$\mathscr{E} = \frac{l^2}{2I},$$

I étant le moment d'inertie du corps par rapport à l'axe de rotation correspondant. La même formule existe en mécanique quantique où la quantité l^2 est quantifiée et ne prend que les valeurs discrètes $l^2 = \hbar^2 l\,(l + 1)$. En mécanique quantique, tout système tournant qui ne change pas est appelé *rotateur*. Par conséquent, les niveaux énergétiques d'un rotateur sont discrets et sont définis par la formule

$$\mathscr{E}_r = \frac{\hbar^2}{2I}\, l\,(l + 1). \tag{31.19}$$

Nous avons déjà utilisé cette formule dans le tome II, § 69 pour faire l'étude qualitative de la capacité calorifique des molécules en rotation.

Il importe de noter que la conception idéalisée du rotateur, tout comme celle d'un solide parfait, est incompatible avec le principe d'incertitude de Heisenberg. En effet, le modèle d'un solide parfait présuppose que ses dimensions suivant une direction quelconque (par exemple suivant l'axe X) sont rigoureusement fixées et ne peuvent changer au cours du temps, ce qui implique que l'incertitude sur la coordonnée $\Delta x = 0$. Mais dans ce cas l'application de la relation d'incertitude implique que suivant la même direction on aurait pour l'impulsion $\Delta p_x = \infty$. Dans ces conditions, le corps serait le siège d'oscillations qui feront varier le moment d'inertie I. Par exemple, dans certains cas, on peut assimiler une molécule en rotation à un rotateur rigide et utiliser la formule (31.19) si les variations de I déterminées par rotation sont petites.

§ 32. Addition des moments angulaires

1. On peut étendre la notion de moment angulaire aux systèmes de particules. On supposera, pour simplifier, que le système se compose de deux particules seulement: *1* et *2*. Notons $\boldsymbol{r}_1$ les coordonnées (le rayon vecteur) de la première particule et $\boldsymbol{r}_2$ celles de la deuxième.

On appelle opérateur moment angulaire $\hat{l}$ du système la somme des opérateurs moments angulaires $\hat{l}_1$, $\hat{l}_2$ de ses parties constitutives:

$$\hat{l} = \hat{l}_1 + \hat{l}_2. \qquad (32.1)$$

On définit de la même façon l'opérateur projection du moment angulaire du système sur une direction donnée. Par exemple,

$$\hat{l}_z = \hat{l}_{1z} + \hat{l}_{2z}. \qquad (32.2)$$

Supposons que les particules n'interagissent pas entre elles. Dans ce cas, la fonction d'onde $\Psi_1(\boldsymbol{r}_1)$ de la première particule sera la même, que la seconde particule soit présente ou absente. On peut multiplier cette fonction par une constante arbitraire qui peut dépendre de $\boldsymbol{r}_2$ comme d'un paramètre. On peut notamment prendre pour constante la fonction d'onde $\Psi_2(\boldsymbol{r}_2)$ de la seconde particule. La fonction d'onde de la première particule se présentera alors sous la forme d'un produit $\Psi = \Psi_1(\boldsymbol{r}_1)\,\Psi_2(\boldsymbol{r}_2)$. En appliquant le même raisonnement à la seconde particule, sa fonction d'onde sera de la même forme. On peut donc considérer $\Psi = \Psi_1\Psi_2$ comme la fonction d'onde du système des particules *1* et *2* n'interagissant pas entre elles. Ce système accepte la normalisation

$$\int |\Psi|^2\, dV_1 dV_2 = \int |\Psi_1|^2\, dV_1 \int |\Psi_2|^2 dV_2 = 1. \qquad (32.3$$

2. Si on se limite à l'action des opérateurs sur la seule fonction $\Psi = \Psi_1\Psi_2$, il résulte de ce qu'on sait déjà que les opérateurs $\hat{l}_1$ et $\hat{l}_2$ sont commutatifs. En effet, puisque l'opérateur $\hat{l}_1$ n'agit que sur la fonction Ψ_1 et que l'opérateur $\hat{l}_2$ n'agit que sur la fonction Ψ_2, on écrira

$$\hat{l}_1\hat{l}_2\Psi = \hat{l}_1\hat{l}_2(\Psi_1\Psi_2) = \hat{l}_1(\Psi_1\hat{l}_2\Psi_2) = (\hat{l}_2\Psi_2)\cdot(\hat{l}_1\Psi_1) = (\hat{l}_1\Psi_1)\cdot(\hat{l}_2\Psi_2).$$

L'action de l'opérateur $\hat{l}_2\hat{l}_1$ conduit au même résultat. Il s'ensui que pour des fonctions de la forme $\Psi = \Psi_1\Psi_2$ est valable l'égalité opératorielle $\hat{l}_1\hat{l}_2 = \hat{l}_2\hat{l}_1$, et c'est ce que démontre notre proposition. Démontrons par le même procédé et pour des fonctions de la même forme la commutativité des opérateurs des projections du moment angulaire (par exemple $\hat{l}_{1x}$ et $\hat{l}_{2x}$) pour un système de particules indépendantes. Il ressort de la démonstration ci-dessus que les règles de commutation (31.6) et toutes les conséquences qui en découlent pour une particule isolée peuvent être appliquées telles quelles aux systèmes de particules indépendantes.

Les opérateurs $\hat{l}_1$ et $\hat{l}_2$ étant commutatifs, l'opérateur du carré du moment angulaire $\hat{l}^2$ sera égal à

$$\hat{l}^2 = (\hat{l}_1 + \hat{l}_2)^2 = \hat{l}_1^2 + 2\,(\hat{l}_1\hat{l}_2) + \hat{l}_2^2. \qquad (32.4)$$

Voyons maintenant comment commute l'opérateur $\hat{l}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2$ avec l'opérateur de l'une de ses projections, par exemple $\hat{l}_x = \hat{l}_{1x} + \hat{l}_{2x}$. De toute évidence, les opérateurs $\hat{l}_{1x}$ et $\hat{l}_{2x}$ commutent avec $\hat{l}_1^2$ et $\hat{l}_2^2$. Il ne reste donc qu'à s'assurer que les opérateurs $\hat{l}_1\hat{l}_2$ et $\hat{l}_x$ sont commutatifs. On a

$$(\hat{l}_1\hat{l}_2)\,\hat{l}_x = (\hat{l}_{1x}\hat{l}_{2x} + \hat{l}_{1y}\hat{l}_{2y} + \hat{l}_{1z}\hat{l}_{2z})\,(\hat{l}_{1x} + \hat{l}_{2x}) =$$
$$= (\hat{l}_{1x}\hat{l}_{2x} + \hat{l}_{1y}\hat{l}_{2y} + \hat{l}_{1z}\hat{l}_{2z})\,\hat{l}_{1x} + (\hat{l}_{1x}\hat{l}_{2x} + \hat{l}_{1y}\hat{l}_{2y} + \hat{l}_{1z}\hat{l}_{2z})\,\hat{l}_{2x},$$

ainsi que

$$\hat{l}_x\,(\hat{l}_1\hat{l}_2) = \hat{l}_{1x}\,(\hat{l}_{1x}\hat{l}_{2x} + \hat{l}_{1y}\hat{l}_{2y} + \hat{l}_{1z}\hat{l}_{2z}) +$$
$$+ \hat{l}_{2x}\,(\hat{l}_{1x}\hat{l}_{2x} + \hat{l}_{1y}\hat{l}_{2y} + \hat{l}_{1z}\hat{l}_{2z}).$$

En appliquant les règles de commutation (31.6) et en mettant à profit le fait que les opérateurs $\hat{l}_1$ et $\hat{l}_2$ agissent sur les fonctions d'onde de particules différentes, on ramène facilement la dernière expression à la précédente. On en conclut que l'opérateur $\hat{l}^2 = (\hat{l}_1 + \hat{l}_2)^2$ commute avec les opérateurs des projections de l sur n'importe quelle direction. Par conséquent, il existe un état du système dans lequel l^2 et une des projections sur un axe quelconque possèdent des valeurs déterminées.

3. On peut caractériser l'état de la première particule par les valeurs l_1 et m_1 et l'état de la deuxième particule par les valeurs l_2 et m_2. Le nombre l_1 détermine le carré du moment angulaire $l_1\,(l_1 + 1)$ de la première particule et l_2 détermine le carré du même moment $l_2\,(l_2 + 1)$ de la seconde particule. Les nombres m_1 et m_2 déterminent les projections sur l'axe Z des moments angulaires l_1 et l_2 (en unités $\hbar$). Il tombe sous le sens que l'ensemble des nombres l_1, l_2, m_1, m_2 caractérise un certain état du système des deux particules indépendantes. La fonction d'onde de cet état sera notée $\varphi_{l_1 l_2 m_1 m_2}$. Evaluons le nombre d'états de ce type, i.e. le nombre de fonctions φ linéairement indépendantes correspondant à des valeurs données de l_1 et l_2. Lorsque l_1 est donné, le nombre m_1 peut prendre $(2l_1 + 1)$ valeurs (cf. § 31, 7) et lorsque l_2 est donné, m_2 peut prendre $(2l_2 + 1)$ valeurs. Ainsi, si l_1 et l_2 sont donnés, le nombre d'états cherché avec les fonctions d'onde du type φ est égal à $(2l_1 + 1)\,(2l_2 + 1)$. Par combinaison linéaire de ces $(2l_1 + 1)\,(2l_2 + 1)$ états, on peut réaliser

n'importe quel état du système caractérisé par des valeurs données de l_1 et l_2.

On peut aussi définir d'une autre manière les états linéairement indépendants à partir desquels on peut réaliser un état caractérisé par des valeurs données de l_1 et l_2. Le nombre total de ces états linéairement indépendants doit rester constant, i.e. il doit toujours être égal à $(2l_1 + 1)(2l_2 + 1)$. Puisque les règles de commutation (31.6) s'appliquent au système tout entier, il doit exister, pour des valeurs données de l_1 et l_2, des états du système tout entier, caractérisés par des valeurs déterminées du carré $l(l + 1)$ du moment angulaire total l et de ses projections m sur l'axe Z. On notera $\psi_{l_1 l_2 l m}$ les fonctions d'onde de ces états. Par combinaison linéaire de ces états, on peut définir la fonction d'onde de n'importe quel état avec l_1 et l_2 donnés. Il s'ensuit que le nombre de fonctions $\psi_{l_1, l_2, l, m}$ linéairement indépendantes et ayant des valeurs données de l_1 et l_2 doit être égal à $(2l_1 + 1)(2l_2 + 1)$.

On peut s'en assurer par un calcul direct. Il découle de la formule (32.2) que si les projections m_1 et m_2 ont des valeurs déterminées, la projection m aura elle aussi une valeur m déterminée telle que $m = m_1 + m_2$. Posons pour fixer les idées que $l_1 > l_2$. Si les valeurs de l_1 et l_2 sont données, les valeurs positives possibles de m, calculées par ce procédé, seront :

m_1	m_2	m
l_1	l_2	l_1+l_2
l_1	l_2-1	l_1+l_2-1
l_1-1	l_2	
l_1	l_2-2	
l_1-1	l_2-1	l_1+l_2-2
l_1-2	l_2	
. . . . . .	. . . . . .	. . . . . .
. . . . . .	. . . . . .	. . . . . .
l_1	$-l_2$	l_1-l_2

Sélectionnons maintenant tous les états possibles (pour l_1 et l_2 donnés) pour lesquels les valeurs maximales de la projection m sont respectivement égales à $(l_1 + l_2)$, $(l_1 + l_2 - 1)$, . . . Ce seront des états caractérisés par une valeur déterminée de l égale à

$$l = (l_1 + l_2),\ l_1 + (l_2 - 1),\ \ldots,\ l_1 - l_2 = l_1 + (l_2 - 2l_2).$$

Le nombre de ces états est égal à $2l_2 + 1$. Dans chacun de ces états, m peut prendre $(2l + 1)$ valeurs. Par conséquent, le nombre d'états possibles (avec l_1 et l_2 donnés) caractérisés par des fonctions

linéairement indépendantes de la forme $\psi_{l_1 l_2 l n}$ sera égal à

$$2(l_1 + l_2) + 1 + 2(l_1 + l_2 - 1) + 1 + \ldots + 2(l_1 - l_2) + 1.$$

C'est une progression arithmétique de raison —2 comportant $2l_2 + 1$ termes. Sa somme est égale à

$$\frac{2(l_1+l_2)+1+2(l_1-l_2)+1}{2}(2l_2+1) = (2l_1+1)(2l_2+1),$$

c.q.f.d.

4. Les résultats obtenus ci-dessus relatifs à l'addition des moments angulaires de deux particules indépendantes se laissent généraliser tels quels à des systèmes complexes quelconques comportant deux parties indépendantes *1* et *2*. Les carrés de leurs moments angulaires (à condition qu'ils aient des valeurs déterminées) sont donnés par les expressions $L_1(L_1 + 1)$ et $L_2(L_2 + 1)$, où L_1 et L_2 sont des entiers positifs. Les projections correspondantes sur l'axe Z (à condition que ces projections aient des valeurs déterminées) peuvent prendre les valeurs suivantes:

$$M_1 = -L_1, -(L_1 - 1), \ldots, +(L_1 - 1), +L_1,$$
$$M_2 = -L_2, -(L_2 - 1), \ldots, +(L_2 - 1), +L_2.$$

Dans ces conditions, le carré du moment angulaire résultant du système tout entier peut prendre les valeurs $L(L + 1)$, où L varie dans les limites

$$L = L_1 + L_2,\ L_1 + L_2 - 1, \ldots, L_1 - L_2, \qquad (32.5)$$

en admettant que $L_1 > L_2$. Les projections correspondantes sur l'axe Z peuvent prendre toutes les valeurs entières comprises entre $M = -L$ et $M = +L$. Ce résultat exprime la *règle d'addition des moments angulaires.*

Dans l'état ainsi défini d'un système complexe, les produits scalaires L_1L_2, LL_1 et LL_2, i.e. les valeurs propres des opérateurs correspondants $\hat{L}_1\hat{L}_2$, $\hat{L}\hat{L}_1$, $\hat{L}\hat{L}_2$, ont aussi des valeurs déterminées. Ce résultat se déduit de la formule (32.4) appliquée aux opérateurs $\hat{L}$, $\hat{L}_1$ et $\hat{L}_2$. Par exemple,

$$\hat{L}_1\hat{L}_2 = \frac{\hat{L}^2 - \hat{L}_1^2 - \hat{L}_2^2}{2}, \qquad (32.6)$$

ou en passant aux valeurs propres

$$L_1L_2 = \frac{L(L+1) - L_1(L_1+1) - L_2(L_2+1)}{2}. \qquad (32.7)$$

On trouve de même

$$LL_1 = \frac{L(L+1) + L_1(L_1+1) - L_2(L_2+1)}{2}, \quad (32.8)$$

$$LL_2 = \frac{L(L+1) + L_2(L_2+1) - L_1(L_1+1)}{2}. \quad (32.9)$$

5. Les résultats qui viennent d'être obtenus sont généralement présentés sous forme de diagrammes vectoriels. Les vecteurs L_1 et L_2 à additionner sont représentés par des segments de droite orientés

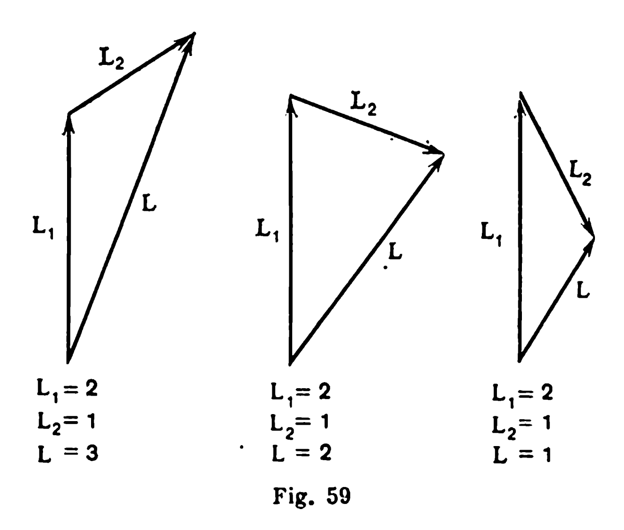


Fig. 59

de modules $\sqrt{L_1(L_1+1)}$ et $\sqrt{L_2(L_2+1)}$ et le vecteur résultant L par un segment orienté de module $\sqrt{L(L+1)}$. A titre d'exemple, la figure 59 représente le diagramme vectoriel pour $L_1 = 2$ et $L_2 = 1$, et pour différents angles entre les vecteurs L_1 et L_2. Trois cas sont à envisager suivant que L prend les valeurs $L_1 + L_2 = 3$, $L_1 + L_2 - 1 = 2$ et $L_1 + L_2 - 2 = 1$. Ce diagramme indique correctement les longueurs de tous les vecteurs, ainsi que leurs produits scalaires, mais il ne reflète pas la vraie nature quantique des moments angulaires, puisque ceux-ci, tout en ayant des longueurs déterminées, n'ont pas d'orientations spatiales.

Si les systèmes *1* et *2* n'interagissent pas entre eux et ne sont pas soumis à l'action de forces extérieures, il y a conservation du moment angulaire résultant L du système tout entier ainsi que des deux moments L_1 et L_2. Si les systèmes *1* et *2* interagissent entre eux, les moments L_1 et L_2 pris séparément ne se conservent pas, mais le moment total L se conserve (en l'absence de forces extérieures). Si l'interaction mutuelle des systèmes est faible, on peut s'attendre à ce

que, comme en mécanique classique, les modules des vecteurs L_1 et L_2 n'en seront pratiquement pas affectés. Sur le diagramme vectoriel les vecteurs L_1 et L_2 exécutent un mouvement de précession, i.e. se mettent à tourner autour du vecteur L avec la même vitesse angulaire. L'étude quantique conséquente de cette question aboutit aux mêmes conclusions.

§ 33. Quantification de l'atome d'hydrogène dans le cas général

1. Dans le § 27, il a été question de la quantification d'un atome d'hydrogène (ou de celle d'un atome hydrogénoïde) sous l'hypothèse que la fonction d'onde ψ possédait une symétrie radiale et ne dépendait donc que de r. Dans ce cas, le moment angulaire de l'électron dans l'atome est égal à zéro, puisque l'opérateur moment n'agit que sur les variables angulaires ϑ et φ, laissant r inchangé. Quant à l'équation de Schrödinger pour les états stationnaires, elle se présente sous la forme

$$\hat{\mathscr{H}}_r \psi = \mathscr{E}\psi, \tag{33.1}$$

l'opérateur $\hat{\mathscr{H}}_r$ étant défini par l'expression

$$\hat{\mathscr{H}}_r = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right) + U(r). \tag{33.2}$$

Cette équation ne décrit que les mouvements quantiques radiaux de l'électron dans l'atome. Il s'agit maintenant de tenir compte de la dépendance de ψ avec les coordonnées angulaires ϑ et φ. Pour ce faire, on doit ajouter à l'opérateur $\hat{\mathscr{H}}_r$ l'opérateur $\frac{1}{2mr^2} l^2$ qui correspond à l'énergie cinétique du mouvement de rotation de l'électron autour du noyau:

$$\hat{\mathscr{H}} = \hat{\mathscr{H}}_r + \frac{1}{2mr^2}\hat{l}^2. \tag{33.3}$$

Il est évident que l'opérateur $\hat{\mathscr{H}}_r$ commute avec les opérateurs $\hat{l}^2$ et $\hat{l}_z$ puisque ces derniers n'agissent que sur les variables angulaires ϑ et φ et sont sans action sur r. La même chose vaut pour l'opérateur $\frac{1}{2mr^2}\hat{l}^2$ puisque la présence du facteur $1/2mr^2$ ne se reflète pas sur la commutativité. Il s'ensuit que l'opérateur $\hat{\mathscr{H}}$ total commute avec $\hat{l}^2$ et l_z et que l'état stationnaire de l'électron dans l'atome d'hydrogène ou dans un atome hydrogénoïde peut être caractérisé par son énergie $\mathscr{E}$, par le carré de son moment angulaire l^2 et par la projection l_z de ce dernier sur l'axe Z.

L'équation de Schrödinger pour les états stationnaires s'écrira maintenant sous la forme

$$\frac{\partial^2\psi}{\partial r^2}+\frac{2}{r}\frac{\partial\psi}{\partial r}+\frac{2m}{\hbar^2}\left(\mathscr{E}-U-\frac{\hat{L}^2}{2mr^2}\right)\psi=0. \tag{33.4}$$

Dans cette expression on recourt à la dérivation partielle par rapport à r puisque la fonction d'onde ψ peut dépendre aussi bien de r que des variables angulaires ϑ et φ. Mais quelle que soit sa dépendance avec ϑ et φ, pour des états stationnaires ayant une valeur déterminée du carré du moment angulaire, on a $\hat{L}^2\psi = L^2\psi = \hbar^2 l(l+1)\psi$. Dans ce cas, l'équation précédente s'écrira

$$\frac{\partial^2\psi}{\partial r^2}+\frac{2}{r}\frac{\partial\psi}{\partial r}+\frac{2m}{\hbar^2}\left(\mathscr{E}-U-\frac{\hbar^2 l(l+1)}{2mr^2}\right)\psi=0. \tag{33.5}$$

Cette équation se distingue de l'équation (33.1) par la présence du terme supplémentaire $-\frac{\hbar^2 l(l+1)}{2mr^2}$. En se plaçant à un point de vue formel, on peut considérer l'équation (33.5) comme la forme de l'équation de Schrödinger dans le cas d'un champ de forces à symétrie radiale avec une fonction de forces potentielle

$$U(r)+\frac{\hbar^2 l(l+1)}{2mr^2}.$$

On peut considérer le deuxième terme de cette somme comme la fonction potentielle de l'électron placé dans le champ d'une force centrifuge et l'équation (33.5) comme l'équation de mouvement de l'électron dans un référentiel animé d'un mouvement de rotation.

Mais puisque les variables indépendantes ϑ et φ ne figurent pas dans l'équation (33.5), la fonction d'onde ψ doit être de la forme $f(\vartheta, \varphi)\,\psi(r)$. Dans tous les calculs la fonction $f(\vartheta, \varphi)$ figurera comme un simple facteur indépendant de r, et se comportera donc comme une constante. Par suite, pour alléger les formules, on omettra partout le facteur $f(\vartheta, \varphi)$ et on fera comme si la fonction ψ ne dépendait que de r. La généralité des résultats que l'on obtiendra n'en sera certainement pas affectée.

2. Jusqu'à présent on n'a pas utilisé la forme explicite de la fonction potentielle $U(r)$. Ayant en vue l'étude de l'atome hydrogénoïde, posons $U(r) = -Ze^2/r$ et introduisons les mêmes notations que dans le § 27, i.e.

$$\beta^2=-\frac{2m\mathscr{E}}{\hbar^2},\quad q=\frac{2mZe^2}{\hbar^2}. \tag{33.6}$$

L'équation (33.5) s'écrira maintenant

$$\frac{\partial^2\psi}{\partial r^2}+\frac{2}{r}\frac{\partial\psi}{\partial r}+\left(\frac{q}{r}-\beta^2-\frac{l(l+1)}{r^2}\right)\psi=0. \tag{33.7}$$

Pour en faire l'étude, on utilisera le même procédé que celui que nous avons mis en œuvre pour l'étude de l'équation (27.1), i.e. on introduira une nouvelle fonction $u(r)$ définie par la relation

$$\psi = \frac{u(r)}{r} e^{-\beta r}. \tag{33.8}$$

Alors

$$\frac{d^2u}{dr^2} - 2\beta \frac{du}{dr} + \left[\frac{q}{r} - \frac{l(l+1)}{r^2}\right] u = 0. \tag{33.9}$$

La solution de cette équation se présente sous forme de la série

$$u = \sum_{k=\gamma}^{\infty} a_k r^k. \tag{33.10}$$

En identifiant les coefficients, on trouve

$$\begin{gathered} \gamma(\gamma - 1) - l(l+1) = 0, \\ [k(k+1) - l(l+1)]\, a_{k+1} = (2\beta k - q)\, a_k \quad \text{pour} \quad k \neq \gamma. \end{gathered} \tag{33.11}$$

La première de ces équations fournit soit $\gamma = (l+1)$, soit $\gamma = -l$. On doit rejeter la valeur $\gamma = -l$ pour les mêmes raisons que celles invoquées au § 27. Il reste donc $\gamma = l + 1$.

Pour étudier la convergence de la série (33.10) on tire de (33.11)

$$\frac{a_{k+1}}{a_k} = \frac{2\beta k - q}{k(k+1) - l(l+1)}.$$

De façon asymptotique

$$\frac{a_{k+1}}{a_k} = \frac{2\beta}{k+1}.$$

Cette expression coïncide en tout point avec l'expression correspondante du § 27, ce qui permet d'en conclure que la série (33.10) présente une coupure. On déduit de la condition de coupure de la série la formule de l'énergie de l'atome (27.8):

$$\mathcal{E} = -\frac{mZ^2e^4}{2\hbar^2n^2}. \tag{33.12}$$

3. Il résulte de l'exposé qui précède que les valeurs de l'énergie dans les états stationnaires de l'atome hydrogénoïde ne dépendent que du nombre quantique principal n. Mais les états avec une valeur donnée de n (i.e. avec une énergie $\mathcal{E}$ donnée) peuvent différer les uns des autres par des valeurs différentes des nombres quantiques l et m. Par conséquent, à une même valeur de $\mathcal{E}$ correspondent plusieurs états quantiques différents. On dit alors que l'état d'énergie $\mathcal{E}$ ou le niveau d'énergie $\mathcal{E}$ est *dégénéré*. Le nombre d'états indépendants dont la superposition permet d'obtenir un état donné d'énergie $\mathcal{E}$ est appelé *degré* ou *multiplicité de dégénérescence*. Déterminons

la multiplicité de dégénérescence pour l'atome hydrogénoïde se trouvant dans un état avec un nombre quantique principal n donné.

Envisageons d'abord les états pour lesquels les deux nombres quantiques n et l ont des valeurs déterminées. Utilisons la formule $\gamma = l + 1$ qui n'a pas encore été utilisée. Comme la série (33.10) doit s'arrêter au terme de n-ième degré, on peut la représenter sous la forme d'une somme finie

$$u = \sum_{k=l+1}^{n} a_k r^k = r^{l+1} \sum_{\alpha=0}^{\alpha=n-l-1} a_{l+\alpha+1} r^{\alpha}. \tag{33.13}$$

Ainsi le nombre quantique principal n est l'exposant majeur dans le polynôme (33.13). Le nombre l est appelé *nombre quantique orbital*. Il caractérise le carré du moment angulaire $l^2 = l(l+1)$ (en unités $\hbar^2$). Ayant fixé la valeur de n, calculons le nombre d'états quantiques ayant des l différents. La plus petite valeur de l est $l = 0$, la plus grande $l = (n - 1)$, puisque dans ce cas la somme (33.13) se réduit à un seul terme. Il s'ensuit que pour une valeur donnée de n le nombre l peut prendre les valeurs suivantes:

$$l = 0,\ 1,\ 2,\ \ldots,\ (n-1), \tag{33.14}$$

i.e. on aura n valeurs en tout et n états quantiques correspondants ayant des valeurs déterminées de n et de l. D'après (33.13) la fonction u peut avoir $n_r = n - l - 1$ nœuds, exclusion faite du nœud $r = 0$. Or, on sait (§ 25, 1) que le nombre de nœuds caractérise le numéro de la fonction d'onde ayant un n donné. On peut donc définir le nombre quantique principal n par la relation

$$n = n_r + l + 1. \tag{33.15}$$

Le nombre quantique n_r fut introduit dans l'ancienne théorie quantique par Sommerfeld et fut appelé *nombre quantique radial*.

Notons que dans un état à l donné, le nombre quantique m (appelé nombre quantique magnétique) caractérisant la projection du moment angulaire sur l'axe Z, peut avoir des valeurs différentes, à savoir:

$$m = -l,\ -(l-1),\ \ldots,\ -1,\ 0,\ +1,\ \ldots,\ +(l-1)\ +l, \tag{33.16}$$

soit au total $2l + 1$ valeurs. Par conséquent, le nombre total d'états quantiques requis pour réaliser un état à n donné, est égal à

$$N = \sum_{l=0}^{l=(n-1)} (2l+1) = n^2. \tag{33.17}$$

Il sera démontré au § 36 qu'en réalité ce nombre doit être doublé pour tenir compte du spin de l'électron. Ainsi, la multiplicité de dégénérescence du niveau énergétique d'un atome hydrogénoïde est égale à $2n^2$.

§ 34. Niveaux énergétiques et séries spectrales des métaux alcalins

1. Dans les atomes des métaux alcalins (lithium, sodium, potassium, rubidium, césium) l'enveloppe électronique contient un électron externe, l'électron dit de valence, qui est assez lachement attaché au noyau. La même structure se retrouve chez les atomes ionisés comportant un seul électron périphérique (électron de valence) (atome d'hélium ionisé une fois, atome de lithium ionisé deux fois, atome de béryllium ionisé trois fois, etc.). Les transitions entre les niveaux énergétiques de l'électron de valence s'accompagnent de l'émission ou de l'absorption de quanta de fréquences relativement basses correspondant à la région optique du spectre. Les autres $Z - 1$ électrons (Z est la charge du noyau exprimée en unités de charge élémentaire) constituent avec le noyau une carcasse relativement stable dans le champ électrique de laquelle se meut l'électron de valence. Les variations de l'énergie des niveaux quantiques de cette carcasse sont relativement grandes et donnent naissance aux spectres de rayons X. On ne s'attardera pas plus sur cette question pour pouvoir consacrer toute notre attention à l'émission et à l'absorption de lumière liées au comportement de l'électron de valence.

Dans ce dernier cas, on peut considérer l'atome du métal alcalin comme un *atome monoélectronique*, où le rôle de noyau est assumé par la carcasse, dont il a été question ci-dessus, et qu'on peut caractériser par une certaine charge efficace $Z_a e$. Pour un atome neutre $Z_a = Z - 1$, pour un atome ionisé une fois $Z_a = Z - 2$, pour un atome ionisé deux fois $Z_a = Z - 3$, etc. Si on éloigne l'électron de valence, la distribution des charges électriques dans la carcasse et son champ électrique acquièrent une symétrie sphérique.

2. Dans ce champ, les états stationnaires de l'électron de valence sont caractérisés par les trois mêmes nombres quantiques que dans l'atome d'hydrogène, à savoir le *nombre quantique principal* n, le *nombre quantique orbital* l et le *nombre quantique magnétique* m. Le nombre quantique l détermine le carré du moment angulaire $l^2 = \hbar^2 l\,(l + 1)$ de l'électron, le nombre magnétique m détermine la projection du moment angulaire sur la direction choisie, généralement confondue avec l'axe Z: $m_z = m\hbar$. Pour une valeur donnée de l, le nombre m peut prendre $2l + 1$ valeurs, à savoir

$$m = -l, -(l - 1), \ldots, -1, 0, +1, \ldots, (l - 1), l.$$

Le nombre quantique principal n est donné par la formule $n = n_r + l + 1$, où n_r est le *nombre quantique radial* qui est égal au nombre de nœuds de la fonction d'onde ψ le long du rayon (le point $r = 0$ n'est pas considéré comme étant un nœud). Pour une valeur donnée de n, le nombre l peut prendre les valeurs suivantes:

$$l = 0, 1, 2, \ldots, (n - 1).$$

Le nombre total d'états quantiques indépendants est donc égal à n^2; à partir de ces états on peut réaliser n'importe quel état ayant un nombre n donné. On verra par la suite qu'on doit associer à ces trois nombres quantiques n, l, m le quatrième nombre quantique, appelé *nombre quantique de spin* m_s, qui peut prendre deux valeurs $m_s = \pm 1/2$. De ce fait, le nombre total d'états quantiques indépendants se trouve doublé. Dans ce paragraphe, on ne tiendra cependant aucun compte du spin de l'électron.

Les différents états de l'électron dans l'atome sont communément dénotés, en fonction de la valeur du nombre l, par des lettres minuscules, conformément au schéma suivant:

Nombre quantique l	0	1	2	3	4	5	6	7
Notation de l'état	s	p	d	f	g	h	i	k

On peut parler d'états s et d'électrons s, d'états p et d'électrons p, etc. Cette terminologie se constitua sous l'influence des premières études spectroscopiques, lorsqu'on n'avait encore aucune notion ni des états quantiques, ni même de la structure de l'atome.

3. Passons à l'examen du comportement de l'électron de valence dans le champ électrique de la carcasse constituée par le noyau et les autres électrons de l'atome. L'électron périphérique exerce une action qui modifie la distribution des charges et le champ électrique de la carcasse. En première approximation, on peut considérer le champ créé par la carcasse comme la superposition du champ produit par la charge ponctuelle $Z_a e$ et du champ d'un dipôle ponctuel, disposés au centre de l'atome. L'axe du dipôle est dirigé vers l'électron périphérique. De ce fait, ce dernier se meut comme si le champ créé par la carcasse, quoique déformé par l'électron, conservait une symétrie sphérique. La fonction potentielle correspondante peut s'écrire sous la forme

$$U = -\frac{Z_a e^2}{r} - C\,\frac{Z_a e^2}{r^2}, \tag{34.1}$$

où C est une constante. On doit considérer cette expression comme les deux premiers termes du développement en série de la fonction U suivant les puissances de $1/r$. Le terme $-C\,\frac{Z_a e^2}{r^2}$ peut être considéré alors comme une correction à apporter au terme principal $-\frac{Z_a e^2}{r}$. Ainsi, dans cette approximation toute la différence d'avec l'atome hydrogénoïde réside en l'apparition du terme $-C\,\frac{Z_a e^2}{r^2}$ dans la fonction potentielle. Dans l'équation (33.5), on peut réunir ce terme avec

l'énergie centrifuge et écrire la somme ainsi obtenue comme suit:

$$\frac{\hbar^2 l(l+1)}{2mr^2} - C\frac{Z_a e^2}{r^2} = \frac{\hbar^2 l^*(l^*+1)}{2mr^2}, \tag{34.2}$$

où le nombre constant l^* est déterminé par l'équation du second degré

$$l^*(l^*+1) = l(l+1) - \frac{2m}{\hbar^2}CZ_a e^2. \tag{34.3}$$

On retombe ainsi sur l'équation (33.7) dans laquelle le nombre l serait remplacé par le nombre l^*, i.e.

$$\frac{\partial^2\psi}{\partial r^2} + \frac{2}{r}\frac{\partial\psi}{\partial r} + \left(\frac{q^*}{r} - \beta^2 - \frac{l^*(l^*+1)}{r^2}\right)\psi = 0, \tag{34.4}$$

avec

$$q^* = \frac{2mZ_a e^2}{\hbar^2}. \tag{34.5}$$

A la différence du nombre l, le nombre l^* n'est pas nécessairement entier, mais cela n'empêche nullement d'appliquer la méthode décrite dans le paragraphe précédent. La fonction $u(r)$ doit être encore cherchée sous la forme de la série (33.10). De (34.3) on trouve pour l^*:

$$l^* = -\frac{1}{2} \pm \sqrt{\left(l+\frac{1}{2}\right)^2 - \frac{2m}{\hbar^2}CZ_a e^2}. \tag{34.6}$$

La quantité positive $\frac{2m}{\hbar^2}CZ_a e^2$ doit être considérée comme la correction à apporter au terme principal $(l + 1/2)^2$. Lorsque cette correction devenait nulle, il fallait prendre le signe plus devant le radical. On doit donc le faire aussi lorsque la correction est différente de zéro. Sans terme correctif $\gamma = l + 1$, tandis qu'avec terme correctif $\gamma = l^* + 1$. Par conséquent, toute la différence entre l'atome d'un métal alcalin et l'atome hydrogénoïde se réduit au remplacement du nombre l par le nombre l^*. Il s'ensuit que les niveaux énergétiques des atomes des métaux alcalins et des ions semblables à ces derniers doivent être déterminés par la formule

$$\mathscr{E} = -\frac{mZ_a^2 e^4}{2\hbar^2(n_r + l^* + 1)^2}.$$

En introduisant le nombre quantique principal (33.15) cette formule devient

$$\mathscr{E} = -\frac{mZ_a^2 e^4}{2\hbar^2(n+\Delta)^2}, \tag{34.7}$$

avec $\Delta = l^* - l$.

4. La formule (34.6) montre que la valeur de la correction Δ dépend du nombre quantique orbital l; la constante C en dépend aussi.

L'origine de cette dernière dépendance se déduit de façon qualitative, mais très clairement, de la théorie de Bohr. Selon cette théorie, à la plus grande valeur de l correspond une orbite circulaire, et à sa plus petite valeur l'orbite elliptique la plus allongée. Dans ce dernier cas, l'orbite de l'électron de valence s'enfonce profondément dans la carcasse de l'atome et perturbe donc fortement son champ électrique. Avec une orbite circulaire il n'y a pas d'interpénétration et la perturbation est donc plus petite. Quoique la mécanique quantique réprouve, par principe, les trajectoires classiques, l'explication qu'elle propose est la même. Du point de vue quantique, la probabilité de présence de l'électron dans les profondeurs de la carcasse atomique est plus grande avec un l petit qu'avec un l grand. En effet, lorsque l est grand, la densité de la probabilité de présence de l'électron dans l'atome est plus proche du cas de la symétrie sphérique que lorsque l est petit. Pour les valeurs maximales $l = n - 1$ elle correspond à la symétrie sphérique, et pour les petites valeurs de l elle s'en éloigne.

Il a été démontré au paragraphe précédent que pour l'atome d'hydrogène ou pour un atome hydrogénoïde à une même valeur de n, à différentes valeurs de l et m correspond un seul et même niveau d'énergie. Par conséquent, l'énergie ne dépend que de n, ce qui signifie qu'il y a *dégénérescence par rapport aux deux nombres quantiques l et m*. Cette dégénérescence est accidentelle et provient de ce que le champ électrique du noyau de l'atome d'hydrogène, i.e. du proton, est un champ coulombien, pour la raison qu'on peut assimiler le proton à une charge ponctuelle. Dans l'atome d'un métal alcalin, l'électron de valence se trouve dans le champ électrique de la carcasse atomique, dont la charge n'est pas ponctuelle quoique sa distribution présente une symétrie sphérique. Comme le champ électrique créé par la carcasse n'est pas inversement proportionnel au carré de la distance au centre de la carcasse, l'énergie de l'électron dépend non seulement du nombre quantique principal n, mais encore du nombre orbital l. Autrement dit, la dégénérescence par rapport à l est levée si le champ qui n'est pas coulombien est à symétrie centrale. La dégénérescence par rapport à m subsiste, car l'énergie ne peut pas dépendre de m par raison de l'isotropie de l'espace. C'est pour cela que les termes spectraux des métaux alcalins diffèrent de ceux de l'atome d'hydrogène.

5. Selon la formule (12.3), aux niveaux d'énergie (34.7) des atomes des métaux alcalins correspondent les termes spectraux

$$T = -\frac{\mathscr{E}}{ch} = -\frac{\mathscr{E}}{2\pi c\hbar} = \frac{Z_a^2 R}{(n+\Delta)^2}, \tag{34.8}$$

où

$$R = \frac{2m\pi^2 e^4}{ch^3} = \frac{me^4}{4c\pi\hbar^3}. \tag{34.9}$$

Cette forme des termes spectraux des métaux alcalins (i.e. pour $Z_a = Z - 1$) a été empiriquement établie par Rydberg (1854-1919) à la fin du XIXe siècle. Ces termes se présentent sous une forme analogue à celle des termes de l'hydrogène, et s'en distinguent par la présence d'une correction Δ, qui est nulle pour l'hydrogène. Il est remarquable que R est précisément la constante de Rydberg figurant dans la formule des termes de l'hydrogène. Ce résultat n'a pas manqué de frapper l'imagination des savants de l'époque.

Les termes spectraux des atomes alcalins sont caractérisés par deux nombres quantiques: le nombre quantique principal n et le nombre quantique orbital l. Pour la notation des termes spectraux, on écrit d'abord le nombre n, sous la forme du chiffre correspondant, puis le nombre l dénoté par une des lettres du tableau donné plus haut. Par exemple $3s$ désigne un terme avec $n = 3$ et $l = 0$; le symbole $5d$ désigne un terme avec $n = 5$ et $l = 2$, etc. Les termes des métaux alcalins sont notés de la façon suivante:

$$\left.\begin{aligned} ns &= \frac{R}{(n+s)^2}, \\ np &= \frac{R}{(n+p)^2}, \\ nd &= \frac{R}{(n+d)^2}, \end{aligned}\right\} \qquad (34.10)$$

etc. Les valeurs numériques de la correction à apporter au nombre n dans le second membre sont dénotées ici par les lettres s, p, d, . . . qu'il ne faut pas confondre avec les mêmes symboles figurant dans le premier membre (ils désignent ici respectivement $l = 0$, $l = 1$, $l = 2$, . . .). Cette terminologie s'est constituée encore du temps des études empiriques des régularités spectrales. On a posé que le nombre Z_a était égal à $Z - 1$, puisqu'il s'agit d'atomes neutres et non pas d'ions.

6. Dans l'atome hydrogénoïde, le niveau énergétique le plus bas est occupé par l'électron à nombre quantique principal $n = 1$. Un autre électron encore peut avoir le même nombre quantique principal (voir § 38). Si on adjoint ce deuxième électron, pour $Z = 2$ on obtient l'atome neutre d'hélium. Dans un atome il ne peut y avoir plus de deux électrons dont le nombre quantique principal n est égal à 1. Ces deux électrons constituent la couche fermée de l'atome d'hélium, ce qu'on a appelé plus haut la carcasse atomique. Si on ajoute un troisième électron, celui-ci ira occuper une nouvelle couche à nombre quantique principal $n = 2$, et pour $Z = 3$ on obtient alors l'élément alcalin — le lithium. Le remplissage de cette couche s'achève avec l'élément néon. Viennent ensuite les métaux alcalins qui commencent la construction des couches fermées suivantes: sodium ($n_{\min} = 3$), potassium ($n_{\min} = 4$), rubidium ($n_{\min} = 5$), césium ($n_{\min} = 6$).

7. Par combinaison des différents termes, on voit apparaître les raies spectrales, conformément au principe de combinaison de Ritz. Une partie seulement des combinaisons possibles des termes, s'accompagnant de l'émission ou de l'absorption de raies spectrales, sont permises. Depuis longtemps, on avait noté, de façon empirique, que dans le cas des métaux alcalins l'émission ou l'absorption de lumière ne se produit que lorsque les termes s se combinent avec les termes p, les termes p avec les termes s et d, les termes d avec les termes p et f, etc. En termes de la théorie atomique, tout cela signifie que dans les transitions quantiques correspondantes, s'accompagnant de l'émission de lumière, le nombre l ne peut varier que d'une unité: $\Delta l = \pm 1$. Ce résultat et des résultats analogues ont reçu le nom de *règles de sélection*. Toutes les autres combinaisons sont *interdites*. Cela ne signifie cependant pas que les transitions correspondantes ne peuvent se produire. Les règles de sélection ne concernent que l'*émission* et l'*absorption de lumière dipolaires*, et ne concernent pas les autres processus pouvant se produire dans les atomes. Par exemple, un choc peut déterminer des transitions du niveau s sur les niveaux d, f, g, etc. Mais comme ces transitions ne modifient pas le moment dipolaire de l'atome, aucune lumière n'est émise. D'ailleurs, les « transitions interdites » peuvent s'accompagner de l'émission de raies spectrales, mais ce ne sera pas un *rayonnement dipolaire* lors duquel le moment dipolaire de l'atome varie, mais un rayonnement *quadru* ou *octupolaire*, résultant d'une variation du moment quadrupolaire ou octupolaire. La règle de sélection $\Delta l = \pm 1$ ne concerne pas ces derniers types de rayonnement. Remarquons encore qu'aucune limitation n'est imposée aux variations du nombre quantique principal n.

La mécanique quantique déchira le voile de mystère qui recouvrait jadis les règles de sélection. Elle formula et résolut un problème plus général: le calcul des *probabilités des transitions* entre les états quantiques des systèmes atomiques s'accompagnant de l'émission de lumière. Il s'avéra que lorsque les règles de sélection n'étaient pas respectées, la probabilité de transition était nulle. On est ainsi conduit aux règles de sélection ou mieux aux *règles d'interdiction* (pour le rayonnement dipolaire de lumière). On peut établir ces règles sans effectuer les calculs des probabilités correspondantes, à l'aide de considérations générales. Il suffit de remarquer que les lois de la Nature présentent une *symétrie par réflexion*, i.e. *sont invariantes par rapport aux opérations d'inversion* (inversion du sens des trois axes de coordonnées). Cette question sera débattue en Physique nucléaire.

8. La règle de sélection $\Delta l = \pm 1$ permet de comprendre pourquoi on doit observer, dans les spectres des métaux alcalins, les *séries spectrales* suivantes (qui furent effectivement observées dans les expériences avant que la théorie ait été élaborée):

Série principale

$$\bar{\nu} = ns - mp. \tag{34.11}$$

Première série latérale (série diffuse)

$$\bar{\nu} = np - md. \tag{34.12}$$

Deuxième série latérale (série nette)

$$\bar{\nu} = np - ms. \tag{34.13}$$

Série de Bergman ou série fondamentale

$$\bar{\nu} = nd - mf. \tag{34.14}$$

Dans ces formules $\bar{\nu} = 1/\lambda$ est le *nombre d'onde spectroscopique*. Dans chacune des séries le nombre n est conservé ; dans la série principale le nombre m parcourt les valeurs $m = n, n + 1, n + 2, \ldots$ et dans les autres séries il parcourt les valeurs $m = n + 1, n + 2, \ldots$, afin qu'apparaissent toutes les raies de chaque série. Dans les limites d'une série, la correction Δ à apporter aux termes variables reste pratiquement constante, mais change d'une série à une autre.

Les figures 60 et 61 représentent les schémas des transitions quantiques et des raies spectrales correspondantes, la première pour le lithium, la seconde pour le sodium. Les longueurs d'onde des raies spectrales sont données en angströms (1Å = 0,1 nm). Sur la figure 60, dans la première colonne (s) sont indiqués les niveaux énergétiques du lithium dans l'état s pour différentes valeurs des nombres quantiques principaux, dans la deuxième les niveaux énergétiques correspondant à différentes valeurs de n et pour des valeurs fixées de p, dans la troisième pour des valeurs fixées de d, etc. Si le champ était coulombien, les niveaux énergétiques dépendraient de n, mais seraient indépendants de l. Pour les métaux alcalins, la dégénérescence des niveaux énergétiques est levée pour la raison que le champ électrique n'est pas coulombien. Les détails du diagramme spectral du sodium ainsi que les questions relatives au symbolisme spectral seront examinés au § 40.

La série principale résulte des transitions quantiques de l'électron de valence des différents états p sur le niveau s le plus profond. Sur ce niveau l'atome se trouve dans un état non excité et peut donc passer sur des niveaux de plus grande énergie. C'est pour cela que la série principale s'observe aussi bien à l'*émission* qu'à l'*absorption*. La figure 62 représente la série principale (partie des courtes longueurs d'onde, à partir de $\lambda = 2594$ Å) du spectre d'absorption de la vapeur de sodium. On y observe la raie de résonance la plus caracté-

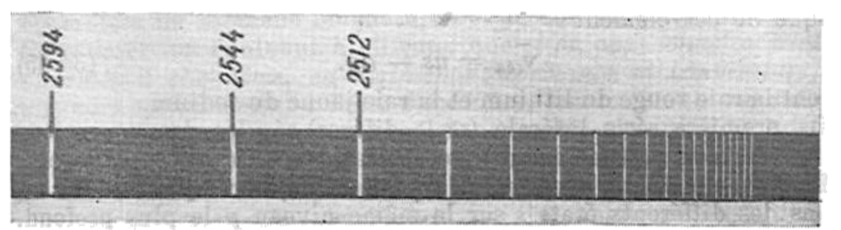


Fig. 62

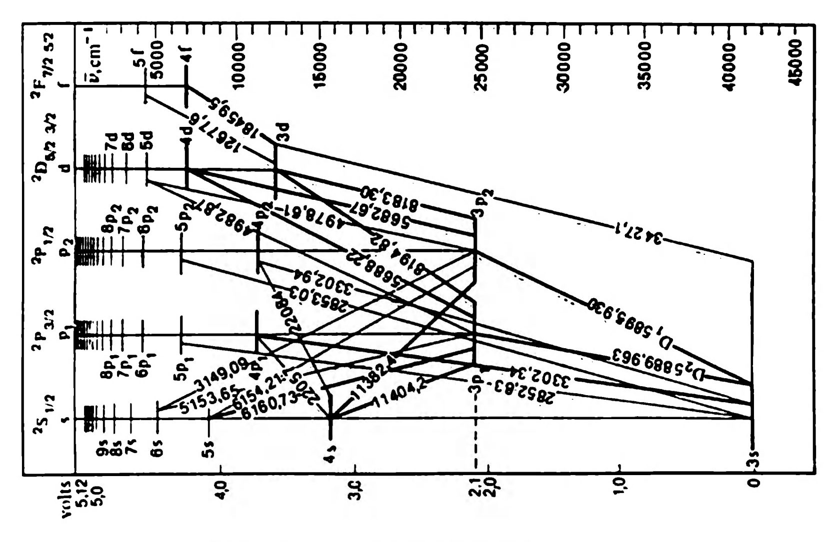


Fig. 61

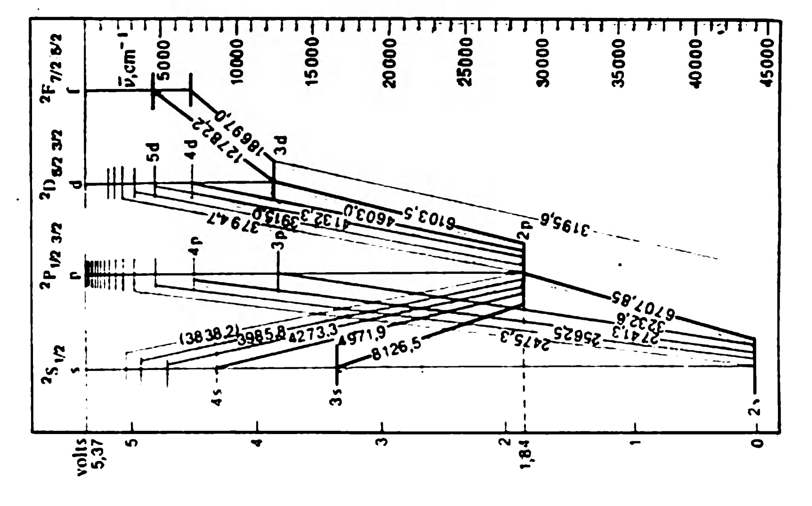


Fig. 60

ristique de cet élément

$$\bar{\nu}_{rés} = ns - np. \qquad (34.15)$$

Ce sont la raie rouge du lithium et la raie jaune du sodium.

La première série latérale (série diffuse) résulte des transitions de l'électron de valence des différents états d sur le niveau p le plus profond ; la deuxième série latérale (série nette) correspond aux transitions des différents états s sur le même niveau p le plus profond. On donnera au § 40 la justification des qualificatifs « diffuse » et « nette ».

PROBLÈMES

1. Les raies spectrales d'atomes dont les noyaux portent la même charge, mais sont de formes et de dimensions différentes, seront-elles, en principe, les mêmes?

Solution. Les spectres seront très légèrement différents, car selon la mécanique quantique, le comportement de l'électron est régi par l'équation d'onde dans tout l'espace où existe un champ de forces.

2. Pour le lithium la valeur du terme $2p$, évaluée à partir des limites des séries latérales, est $2p = 28581{,}4$ cm^{-1}. Les longueurs d'onde des raies $2p - 3d$ et $3d - 4f$ sont respectivement égales à $\lambda_1 = 6103{,}77$ Å et $\lambda_2 = 18697{,}0$ Å. Calculer la longueur d'onde de la raie $2p - 4f$.

Solution. Les nombres d'onde spectroscopiques des raies $2p - 3d$ et $3d - 4f$ sont respectivement égaux à

$$1/\lambda_1 = \bar{\nu}_1 = 16383{,}3 \text{ cm}^{-1},$$

$$1/\lambda_2 = \bar{\nu}_2 = 5348{,}4 \text{ cm}^{-1}.$$

D'où $3d = 2p - \bar{\nu}_1 = 12198{,}1$ cm^{-1}, $4f = 3d - \bar{\nu}_2 = 6855{,}5$ cm^{-1}; $2p - 4f = 21725{,}9$ cm^{-1}, $\lambda = 4602{,}8$ Å.

§ 35. Le magnétisme des atomes

1. Depuis Ampère (1775-1836) le magnétisme a été réduit aux *courants électriques* qui, selon les idées d'Ampère, circulent dans les particules ultimes (atomes et molécules) de la matière. La nature physique de ces courants a pû être élucidée dès l'apparition des conceptions de la structure électronique de la matière et de la théorie de Bohr. On admettait que les *courants moléculaires postulés par Ampère* étaient créés par les électrons tournant autour du noyau de l'atome. Mais la physique classique ne pouvait, avant que furent énoncées les conceptions quantiques, expliquer ni le mouvement des électrons autour du noyau, ni l'existence même des atomes. La mise en œuvre des méthodes de la physique statistique permit de démontrer que *du point de vue classique dans l'état stabilisé la substance ne pouvait être aimantée, i.e. ne pouvait avoir un moment magnétique différent de zéro* (Bohr, Lorentz, Van Leeuwen, cf. t. III, § 75). Cela ne signifie nullement qu'il soit impossible d'aimanter les substances, puisqu'on peut mettre en rotation les charges électriques, i.e. y exciter des courants circulaires. Dans ces conditions, on verrait

apparaître un moment magnétique et la substance serait aimantée. Cette assertion équivaut à affirmer que *si on abandonne la substance aimantée à elle-même, en maintenant constante sa température, elle atteindra spontanément à l'état d'équilibre, excluant toute aimantation, même si la substance se trouve placée dans un champ magnétique extérieur.* Or, cette conclusion est en désaccord avec les faits expérimentaux.

On n'arriva à comprendre la nature du magnétisme qu'après l'avènement de la mécanique quantique. Le magnétisme, tout comme l'existence des atomes et des molécules, sont des *effets quantiques*. Les théories classiques de l'aimantation (Langevin) ont remporté des succès notables uniquement parce qu'elles introduisaient tacitement des hypothèses de *caractère essentiellement quantique*, à savoir l'existence chez les atomes de moments magnétiques préexistants ou d'orbites stationnaires sur lesquelles se mouvaient les électrons. Or, c'est la théorie qui devrait expliquer ces hypothèses.

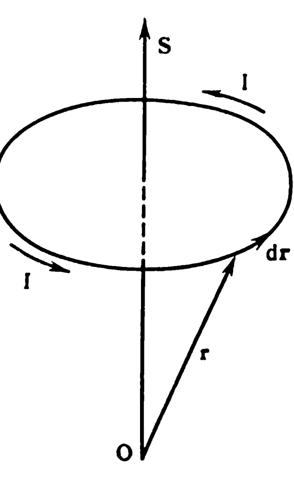


Fig. 63

2. Puisque les électrons constituant l'enveloppe des atomes portent des charges et possèdent une masse, à leur mouvement (appelé mouvement orbital) au sein de l'enveloppe doit être associé non seulement un moment cinétique, mais aussi un *moment magnétique atomique*. Dans le tome III (§ 75) on a déjà examiné le lien existant entre ces moments, dans la mesure où on pouvait le faire sans faire appel aux conceptions quantiques. Le lien qui y a été établi reste valable en mécanique quantique, mais sa signification et, partant, sa justification diffèrent de celles de la mécanique classique, puisque le concept de moment cinétique ne peut être transposé directement de la théorie classique dans la théorie quantique. On doit pour cela introduire l'opérateur correspondant. La même chose vaut pour la notion de moment magnétique. On prendra pour point de départ la théorie classique.

Selon l'électrodynamique (cf. t. III, § 57) une spire circulaire parcourue par un courant continu I (fig. 63) possède un moment magnétique égal à

$$\vec{\mathfrak{m}} = \frac{I}{c} S, \qquad (35.1)$$

où S est le vecteur de la surface limitée par le contour de courant. Ce vecteur est défini par la formule

$$S = \frac{1}{2} \oint [r\, dr]$$

et ne dépend pas du choix de l'origine O des coordonnées puisque le contour de courant est fermé. On supposera que le sens de parcours du contour coïncide avec le sens du courant. Le sens de parcours du contour et le sens du vecteur S sont dans le rapport correspondant à un trièdre de référence direct. Le moment magnétique d'un contour fermé de courant continu peut être représenté sous la forme

$$\vec{\mathfrak{m}} = \frac{1}{2c} \oint I\,[\boldsymbol{r}\, d\boldsymbol{r}]. \qquad (35.2)$$

Comme le courant I résulte du mouvement des charges, ce sont les charges mobiles qui créent le moment magnétique $\vec{\mathfrak{m}}$. Chaque charge en mouvement crée un moment magnétique. Le moment magnétique total d'un corps s'obtient en faisant la somme vectorielle des moments magnétiques produits par les charges mobiles qu'il contient. Il convient donc de transformer l'intégrale de contour (35.2) en une intégrale étendue à toutes les charges mobiles du corps. Soit dq la charge qui traverse dans l'unité de temps dt la section droite de la spire parcourue par le courant (s'il s'agit d'un courant continu, cette quantité ne dépend pas de la position de la section droite considérée). On a alors $I = dq/dt$,

$$\vec{\mathfrak{m}} = \frac{1}{2c} \oint \frac{dq}{dt}\,[\boldsymbol{r}\, d\boldsymbol{r}] = \frac{1}{2c} \oint \left[\boldsymbol{r}\, \frac{d\boldsymbol{r}}{dt}\right] dq.$$

Comme dans cette formule on intègre aussi sur $d\boldsymbol{r}$, l'intégrale reste de contour. Choisissons l'élément de contour $d\boldsymbol{r}$ de telle sorte que, dans le temps dt, la charge dq se déplace de $d\boldsymbol{r}$. On aura alors $d\boldsymbol{r} = \boldsymbol{v}\, dt$ et on obtient

$$\vec{\mathfrak{m}} = \frac{1}{2c} \int [\boldsymbol{r}\boldsymbol{v}]\, dq = \frac{1}{2\mu c} \int [\boldsymbol{r}\boldsymbol{p}]\, dq, \qquad (35.3)$$

où $\boldsymbol{v}$ est la vitesse, $\boldsymbol{p}$ l'impulsion et μ la masse liée à la charge mobile dq. (On dénote la masse par la lettre μ parce que m dénote le nombre quantique magnétique.)

Dans ces conditions, dq représente la charge contenue, à l'instant considéré, dans l'élément $d\boldsymbol{r}$ du contour; le temps dt se trouve ainsi éliminé de la formule (35.3). Cette formule ne se rapporte plus à une spire parcourue par un courant continu (le petit cercle sur le symbole de l'intégrale peut donc être omis). Ne subsiste qu'un système de charges, où chaque charge est caractérisée, outre sa grandeur, par sa position et sa vitesse de déplacement. Or ce sont les seules grandeurs qui importent pour la création d'un moment magnétique. Il n'importe nullement de savoir comment s'est consitué le système de charges et quel est son état.

La formule (35.3) représente le moment magnétique d'un corps comme la superposition des moments magnétiques des charges mo-

biles. En généralisant, on peut la récrire sous la forme

$$\vec{\mathfrak{m}} = \frac{1}{2c} \sum_i q_i [r_i v_i]. \qquad (35.4)$$

Cette formule suppose qu'il s'agit d'un système de charges ponctuelles q_i animées à l'instant considéré de vitesses v_i. On ne fait aucune hypothèse sur la nature du mouvement.

3. L'expression classique (35.3) du moment magnétique d'un système de charges mobiles dépend du choix de l'origine des coordonnées. En effet, si a est le rayon vecteur de la nouvelle origine (lettres affectées d'accents) par rapport à l'ancienne (lettres sans accents), on aura pour toutes les charges $r_i = r_i' + a$, de sorte que

$$\vec{\mathfrak{m}} = \vec{\mathfrak{m}}' + \frac{1}{2c} \sum_i q_i [a v_i].$$

Il s'ensuit que l'ancien moment magnétique $\vec{\mathfrak{m}}$ et le nouveau $\vec{\mathfrak{m}}'$ ne seront toujours égaux que si pour tout vecteur a le produit vectoriel $[a \sum q_i v_i]$ est nul. Cela se réalise notamment pour toute spire fermée et immobile, parcourue par un courant continu, puisque alors $\sum q_i v_i = 0$.

4. Pour une seule charge ponctuelle animée d'une vitesse v

$$\vec{\mathfrak{m}} = \frac{q}{2c} [rv] = \frac{q}{2\mu c} [rp], \qquad (35.5)$$

où μ est la masse et p l'impulsion de la particule portant cette charge. Ainsi, la physique classique aboutit à la relation

$$\vec{\mathfrak{m}} = \Gamma L, \qquad (35.6)$$

où

$$\Gamma = \frac{q}{2\mu c}. \qquad (35.7)$$

Ces formules avaient déjà été établies dans le tome III, § 75 par un raisonnement plus simpliste. Pour l'électron $q = -e$, et donc

$$\Gamma = -\frac{e}{2\mu_e c}. \qquad (35.8)$$

Dans ce cas le rapport Γ du moment magnétique de l'électron au moment mécanique est appelé *rapport gyromagnétique* du mouvement orbital de l'électron.

Notons que toutes les relations ci-dessus ont été établies dans le cadre de la *mécanique non relativiste* (on n'a pas tenu compte de la variation de la masse avec la vitesse et de ce que les particules étaient considérées comme ponctuelles). D'ailleurs cette dernière circonstance n'est pas essentielle et les particules peuvent avoir de l'étendue,

puisque alors on peut les diviser mentalement en des parties assez petites pour pouvoir les assimiler à des points. Mais pour que le rapport $\mathfrak{m}/L$ ne change pas, il faut supposer que la distribution spatiale des charges et des masses obéit à une même loi. Pour une bille chargée tournant autour d'un de ses diamètres avec une vitesse non relativiste, la physique classique conduit aux formules (35.7) et (35.8) indépendamment du mode de distribution des masses et des charges : la seule chose qui importe est que leurs distributions soient les *mêmes*. Bien entendu, il en ira tout autrement si, par exemple, la charge est concentrée au centre de la bille, tandis que la masse est uniformément répartie dans tout le volume de la bille.

5. Il convient de remplacer maintenant les conceptions classiques par des conceptions quantiques. En mécanique quantique la formule (35.5) ne peut servir de définition du moment magnétique puisqu'il n'existe aucun état d'une particule qui pourrait être caractérisé par sa position exacte $\boldsymbol{r}$ et par son impulsion exacte $\boldsymbol{p}$. De même que dans le cas du moment angulaire, la mécanique quantique remplace la formule classique (35.5) par la *relation opératorielle*

$$\hat{\vec{\mathfrak{m}}} = \frac{q}{2c}[\hat{\boldsymbol{r}}\hat{\boldsymbol{v}}] = \frac{q}{2\mu c}[\hat{\boldsymbol{r}}\hat{\boldsymbol{p}}], \qquad (35.9)$$

ou

$$\hat{\vec{\mathfrak{m}}} = \Gamma\hat{\boldsymbol{L}}. \qquad (35.10)$$

L'étude du moment magnétique d'une particule se ramène donc à l'étude des propriétés de l'opérateur $\hat{\vec{\mathfrak{m}}}$. Puisque les opérateurs $\hat{\vec{\mathfrak{m}}}$ et $\hat{\boldsymbol{L}}$ ne diffèrent que par un facteur constant, leurs propriétés doivent être parfaitement semblables. Notons, par exemple, que l'opérateur $\hat{\vec{\mathfrak{m}}}$, tout comme $\hat{\boldsymbol{L}}$, *ne dépend pas du choix de l'origine des coordonnées.* La quantification des moments magnétique et angulaire se fait suivant les mêmes règles. Les composantes du moment magnétique sur deux directions différentes *ne peuvent avoir simultanément des valeurs déterminées.* A l'état stationnaire, ne peuvent avoir des valeurs déterminées que le carré du moment magnétique et l'une de ses projections sur un des axes de coordonnées, généralement l'axe Z. On déduit directement des formules (35.8) et (35.10) relatives au mouvement orbital de l'électron le résultat suivant :

$$\mathfrak{m}_z = -\frac{e}{2\mu_e c}L_z = -\mathfrak{m}_B m, \qquad (35.11)$$

où

$$\mathfrak{m}_B = \frac{c\hbar}{2\mu_e c} = 9{,}273\cdot 10^{-21}\ \text{erg}\cdot\text{Gs}^{-1}. \qquad (35.12)$$

La constante $\mathfrak{m}_B$ est appelée *magnéton de Bohr.* On peut considérer le magnéton de Bohr comme le *quantum de moment magnétique* (plus exactement de sa projection sur une direction donnée).

On établit parfois la formule (35.11) d'une façon différente. Partant de l'équation temporelle de Schrödinger, on obtient l'équation de continuité $\partial\rho/\partial t + \operatorname{div} j = 0$, où ρ et j sont respectivement la densité de probabilité et la densité du courant de probabilité. Connaissant la valeur de cette dernière quantité et de la fonction d'onde, on trouve la densité de probabilité du courant électrique dans l'état stationnaire de l'atome; en intégrant directement, on trouve également le moment magnétique moyen créé par ce courant. Ce raisonnement direct possède un défaut de principe, à savoir que la densité du courant de probabilité j déterminée par l'équation de Schrödinger non relativiste *n'est pas univoque*, puisqu'on peut ajouter à l'expression obtenue un terme quelconque de la forme rot $\boldsymbol{a}$ (puisque div rot $\boldsymbol{a} \equiv 0$) sans affecter le courant de probabilité total passant par une surface fermée, qui est la seule grandeur observable. La densité du courant électrique circulant dans l'atome est, à la différence du courant de probabilité, une *grandeur observable*, mais pour la déterminer de façon univoque *la seule équation de Schrödinger non relativiste ne peut suffire*. Pour lever la non-univocité il faut faire appel à la *théorie relativiste*. En effet, de par son essence, ρ est une quantité *univoque* et en théorie relativiste le scalaire ρ et les trois composantes du vecteur j forment un *quadrivecteur relativiste invariant*, dont la composante temporelle est ρ.

§ 36. Les expériences de Stern et Gerlach. Le spin de l'électron

1. En 1921, Stern (1888-1969) et Gerlach (1889-1979) ont démontré par des expériences directes que les atomes possèdent des moments magnétiques et que ces moments magnétiques sont quantifiés. Par vaporisation, dans le four K, de l'élément étudié, on obtenait dans un vide poussé un jet d'atomes nettement délimité à l'aide des diaphragmes B et B' (fig. 64). Le jet atomique traversait un intense champ magnétique H régnant entre les pièces polaires N et S d'un électroaimant. L'une des pièces polaires (N) était taillée en forme de prisme à arête aigüe, tandis qu'un canal était creusé dans l'autre pièce polaire (S). On obtenait ainsi un champ magnétique *très inhomogène*. Après traversée du champ magnétique, le jet atomique tombait sur la plaque photographique P et y laissait sa trace.

Procédons à un calcul du comportement du jet d'atomes en nous plaçant d'abord au point de vue classique et en supposant donc que les moments magnétiques ne sont pas quantifiés. Si $\vec{\mathfrak{m}}$ est le moment magnétique de l'atome, ce dernier est soumis dans le champ magnétique inhomogène à l'action d'une force

$$f = (\vec{\mathfrak{m}}\, \nabla)\, \boldsymbol{H}.$$

Orientons l'axe Z suivant le sens du champ magnétique (dans le sens allant de N à S normalement aux pièces polaires). La projection de la force dans cette direction est alors égale à

$$f_z = \mathfrak{m}_x \frac{\partial H_z}{\partial x} + \mathfrak{m}_y \frac{\partial H_z}{\partial y} + \mathfrak{m}_z \frac{\partial H_z}{\partial z} .$$

Les deux premiers termes de cette expression ne jouent aucun rôle, puisque selon les conceptions classiques, un atome placé dans

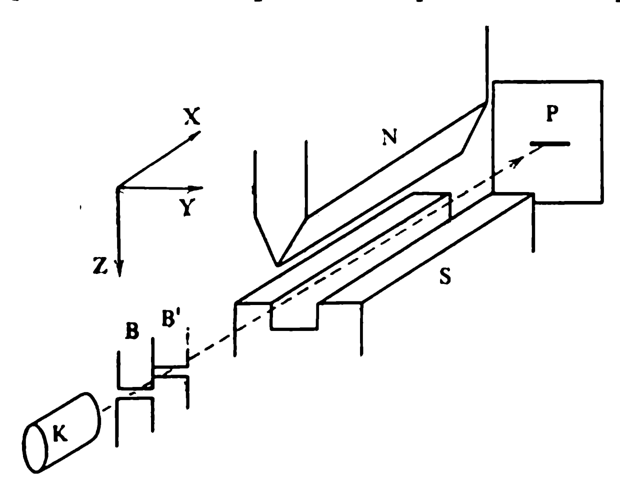


Fig. 64

un champ magnétique y exécute un mouvement de précession autour de l'axe Z, avec la fréquence de Larmor

$$\Omega = -\frac{eH}{2\mu c}$$

(la charge de l'électron est notée $-e$). Il s'ensuit que les projections $\mathfrak{m}_x$ et $\mathfrak{m}_y$ oscillent avec la même fréquence et prennent alternativement des valeurs positives et négatives. Si la vitesse angulaire de la précession est suffisamment grande, on peut prendre la moyenne temporelle de f_z. Dans ce cas, les deux premiers termes de l'expression de f_z s'annulent et on peut écrire

$$\bar{f}_z = \mathfrak{m}_z \frac{\partial H_z}{\partial z}. \qquad (36.1)$$

Pour nous faire une idée à quel point ce moyennage est admissible, procédons à une estimation numérique. La période de la précession de Larmor est

$$\tau = \frac{2\pi}{|\Omega|} = \frac{4\pi\mu c}{eH} \approx \frac{7\cdot 10^{-7}}{H} \text{ s},$$

le champ H étant donné en gauss. Si, par exemple, $H = 1000$ Gs

on trouve $\tau \approx 7 \cdot 10^{-10}$ s. Si dans le jet la vitesse des atomes est égale à $v = 100$ m/s $= 10^4$ cm/s, dans le temps τ un atome parcourt la distance $x \approx 7 \cdot 10^{-6} \approx 10^{-5}$ cm, i.e. une distance négligeable devant les dimensions caractéristiques du dispositif expérimental. C'est ce qui démontre que le moyennage de f_z sur le temps est admissible.

Mais la mécanique quantique permet de justifier également la formule (36.1). En effet, l'application d'un fort champ magnétique suivant l'axe Z fait passer l'atome dans un état caractérisé par une seule composante déterminée du moment magnétique, à savoir $\mathfrak{m}_z$. Dans cet état les deux autres composantes $\mathfrak{m}_x$ et $\mathfrak{m}_y$ ne peuvent avoir de valeurs déterminées. Si on mesurait $\mathfrak{m}_x$ et $\mathfrak{m}_y$ dans cet état, on obtiendrait des valeurs différentes dont la moyenne serait nulle. Il s'ensuit que dans le cas d'une étude quantique le moyennage (36.1) est justifié.

2. Néanmoins, on doit s'attendre à ce que les prévisions de la théorie classique et de la théorie quantique soient différentes quant aux résultats de l'expérience de Stern et Gerlach. Dans ces expériences, on enregistrait la trace du jet atomique d'abord pour le champ magnétique nul, puis pour le champ branché. Si la projection $\mathfrak{m}_z$ pouvait prendre n'importe quelles valeurs continues, comme l'impose la théorie classique, la force $\overline{f_z}$ devrait également prendre des valeurs continues possibles et l'application du champ n'aurait provoqué qu'un *élargissement du jet atomique*. Les prévisions de la théorie quantique sont très différentes. Dans ce cas, la projection $\mathfrak{m}_z$ et par suite la force moyenne $\overline{f_z}$ sont *quantifiées*, autrement dit ne peuvent avoir que certaines valeurs discrètes. Si le nombre quantique orbital de l'atome est égal à l, la théorie prévoit que le jet sera scindé en $2l + 1$ jets (i.e. le nombre de jets est égal au nombre de valeurs possibles du nombre quantique m). Par conséquent, suivant la valeur de l, on doit s'attendre à ce que *le jet initial soit scindé en* 1, 3, 5, . . . *composantes*. Le nombre des composantes doit toujours être *impair*.

Les expériences de Stern et Gerlach démontrèrent la quantification de la projection $\mathfrak{m}_z$, mais leurs résultats n'étaient pas toujours conformes à la théorie qui vient d'être exposée. Dans les premières expériences ils utilisèrent des jets d'atomes d'argent. Dans le champ magnétique le jet initial se scindait en deux composantes. La même chose se produisait avec l'hydrogène atomique, mais avec les atomes d'autres éléments le schéma de la scission était non seulement plus compliqué, mais le nombre des composantes pouvait être *impair*, comme le veut la théorie, ou *pair* en contradiction avec les prévisions de celle-ci. Il fallait donc corriger la théorie.

3. A tout cela viennent s'ajouter les résultats des expériences d'Einstein et de Haas (1878-1966) et de celles de Barnett (1873-1956) relatifs à la détermination du rapport gyromagnétique et qui ont

été décrits dans le tome III (§ 78). On trouva par exemple que pour le fer le rapport gyromagnétique était égal à

$$\frac{\mathfrak{m}}{L}=\left|\frac{e}{\mu_e c}\right|, \tag{36.2}$$

donc deux fois plus grand que ne le prévoyait la théorie.

On découvrit encore que les termes spectraux des métaux alcalins présentaient une *structure* dite *en doublet*, i.e. se composaient de deux niveaux rapprochés. Pour décrire cette structure les trois nombres quantiques n, l, m ne suffisaient plus et il fallut introduire un *quatrième nombre quantique*. C'était surtout pour cela que Uhlenbeck (né en 1900) et Goudsmit (1902-1979) formulèrent en 1925 l'hypothèse du *spin de l'électron* *). Selon cette hypothèse, l'électron posséderait, outre le moment cinétique et le moment magnétique, tous deux liés au déplacement en bloc de cette particule, un *moment cinétique propre* ou *interne*, et se comporterait donc à ce point de vue comme une toupie classique. Ce moment cinétique propre est appelé *spin* (du verbe anglais « to spin » — tourner sur soi-même). Le moment magnétique correspondant est appelé *moment magnétique de spin*. Ces moments sont dénotés respectivement par L_s et $\vec{\mathfrak{m}}_s$ pour les distinguer des moments orbitaux L_l et $\vec{\mathfrak{m}}_l$. Le plus souvent on dénote le spin par la lettre s.

Dans les expériences de Stern et Gerlach les atomes d'hydrogène se trouvaient dans l'état s et ne possédaient donc pas de moments orbitaux. Comme le moment magnétique du noyau est extrêmement petit, Uhlenbeck et Goudsmit supposèrent que la scission du jet atomique était due au *moment magnétique de spin* et non pas au moment orbital. La même conclusion a été faite dans le cas d'atomes d'argent. L'atome d'argent ne possède qu'un seul électron périphérique, et sa carcasse étant symétrique ne possède ni moment de spin, ni moment magnétique. Tout le moment magnétique de l'atome d'argent est créé par le seul électron périphérique. Lorsque l'atome se trouve dans son état normal, i.e. dans l'état s, le moment orbital de l'électron de valence est nul et tout le moment de l'atome est dû au spin de cet électron.

4. Uhlenbeck et Goudsmit supposaient que le spin résulte de la rotation de l'électron autour de son axe. Le modèle planétaire de l'atome qui était admis à l'époque ressembla encore plus au système solaire. Les électrons (les planètes) tournent non seulement autour du noyau (du Soleil) mais encore autour de leurs propres axes. Mais tout aussitôt l'inadéquation de cette conception classique du spin devient évidente (voir problème n° 1 à la fin du paragraphe). Pauli introduisit le spin de façon systématique dans la mécanique quanti-

*) L'hypothèse du spin électronique fut suggérée par Kronig (né en 1904), mais sa note ne fut pas publiée.

que, mais exclua toute possibilité de son interprétation classique. En 1928, Dirac montra que le spin de l'électron fait partie de la théorie de l'électron fondée sur l'équation d'onde relativiste. La théorie de Dirac fait intervenir aussi le moment magnétique de spin de l'électron, et on y trouve pour le rapport gyromagnétique une valeur conforme aux résultats expérimentaux. Rien ne fut dit sur la structure interne de l'électron qui était considéré comme une particule ponctuelle possédant une masse et une charge. Ainsi, *le spin de l'électron se révéla être un effet quantique relativiste ne pouvant être interprété par la théorie classique.* Plus tard le concept de spin en tant que moment cinétique propre fut étendu à d'autres particules élémentaires ou composées et, ayant été prouvé, trouva de nombreuses applications dans la physique moderne.

5. Il est bien évident qu'on ne peut développer, dans un cours de physique générale, la théorie complète du spin. On posera, en qualité de proposition de base, qu'au spin s on peut faire correspondre un opérateur vectoriel $\hat{\mathbf{s}}$ dont les projections $\hat{s}_x$, $\hat{s}_y$, $\hat{s}_z$ vérifient, au même titre que les projections de l'opérateur moment orbital, les règles de commutation (31.2), i.e.

$$\begin{aligned} \hat{s}_y \hat{s}_z - \hat{s}_z \hat{s}_y &= i\hbar \hat{s}_x, \\ \hat{s}_z \hat{s}_x - \hat{s}_x \hat{s}_z &= i\hbar \hat{s}_y, \\ \hat{s}_x \hat{s}_y - \hat{s}_y \hat{s}_x &= i\hbar \hat{s}_z. \end{aligned} \tag{36.3}$$

Il résulte de ces relations que dans un état donné ne peuvent avoir de valeur déterminée que le carré du spin total $\mathbf{s}^2$ et l'une de ses projections sur un axe donné (que l'on fait généralement coïncider avec l'axe Z). Si la plus grande valeur de la projection s_z (en unités $\hbar$) est égale à s, le nombre de toutes les projections possibles correspondant à s donné sera égal à $2s + 1$. Les expériences de Stern et Gerlach ont montré que pour l'électron ce nombre était égal à 2, i.e. $2s + 1 = 2$, d'où $s = 1/2$. La valeur maximale (en unités $\hbar$) que peut prendre la projection du spin sur une direction donnée, i.e. le nombre s, représente la valeur du spin de la particule.

Le spin des particules peut être entier ou demi-entier. Pour l'électron le spin est égal à 1/2. Il résulte des relations (36.3) que le carré du spin d'une particule est $\mathbf{s}^2 = s(s + 1)$, ce qui donne pour l'électron $\mathbf{s}^2 = (1/2)(1/2 + 1) = 3/4$ (en unités $\hbar^2$).

Des mesures de la projection $\mathfrak{m}_z$ du moment magnétique, réalisées par la méthode de Stern et Gerlach, ont montré que pour les atomes d'argent et d'hydrogène la valeur de $\mathfrak{m}_z$ était égale au magnéton de Bohr $\mathfrak{m}_B$ (formule 35.12). Ainsi, le rapport gyromagnétique de l'électron est $\left|\frac{\mathfrak{m}_z}{L_z}\right| = \frac{\mathfrak{m}_B}{{}^1/{}_2\hbar} = \frac{e}{\mu_e c}$,

en accord avec la formule (36.2).

PROBLÈMES

1. En se plaçant au point de vue de la théorie classique, on a essayé d'interpréter le spin électronique comme un moment cinétique apparaissant du fait que l'électron tourne sur soi-même autour d'un diamètre. On assimilait l'électron à une bille et on calculait sa masse par la formule relativiste $\mathscr{E} = \mu c^2$; on postulait que l'énergie propre de l'électron était d'origine purement électrostatique. Analyser les difficultés que fait apparaître une telle interprétation classique du spin.

S o l u t i o n. Lorsque l'électron tourne sur lui-même, son moment cinétique est égal à $L = {}^1/_2\hbar = I\omega$, I étant le moment d'inertie de l'électron par rapport à son diamètre et ω la vitesse angulaire. De toute évidence $I < \mu r^2$, puisque pour atteindre la valeur maximale de I il faudrait que toute la masse de l'électron soit concentrée sur l'équateur de l'électron en rotation. Par conséquent,

$$^1/_2\hbar < \mu r^2\omega,$$

d'où

$$v > \frac{\hbar}{2\mu r},$$

où $v = \omega r$ est la plus grande vitesse linéaire à la surface de l'électron. La masse de l'électron $\mu = \mathscr{E}/c^2$. Lorsqu'on cherche à déterminer l'énergie propre $\mathscr{E}$ de l'électron, on se heurte à la difficulté suivante. La valeur de $\mathscr{E}$ dépend de la distribution de la charge totale e de l'électron dans son volume. Le minimum de $\mathscr{E} = e^2/2r$ doit s'obtenir lorsque la charge e serait répartie sur la surface de l'électron. Si la charge était répartie uniformément dans son volume, on aurait $\mathscr{E} = {}^3/_5 e^2/r$. En posant $\mathscr{E} = e^2/r$, on trouve

$$v = \frac{\hbar}{2\mathscr{E}r/c^2} > \frac{\hbar c^2}{2e^2} = \frac{c}{2\alpha},$$

où $\alpha = e^2/\hbar c$ est une quantité sans dimension appelée *constante de la structure fine* ($\alpha \approx 1/137$). Ainsi, avec les hypothèses faites, on trouve $v > 68,5\,c$; comme v est plus grand que la vitesse de la lumière c, ce résultat est absurde.

2. Démontrer que par la méthode de Stern et Gerlach on ne peut pas mesurer le moment magnétique de l'électron si on opère dans l'expérience avec un *électron libre* et non pas avec un électron lié à l'atome.

S o l u t i o n. Les dimensions transversales Δx du pinceau de particules dans toutes les directions doivent satisfaire à la condition $\Delta x \gg \lambda$, où λ est la longueur de l'onde de de Broglie pour les particules considérées, sinon le pinceau s'étalera rapidement par suite de la diffraction. Dans les expériences de Stern et Gerlach le pinceau d'atomes (plus exactement le pinceau d'ions de charge e) est dévié (scindé) suivant la direction de l'axe Z (fig. 64). La force moyenne déterminant cette déviation est donnée par la formule (36.1). Compte tenu de l'équation $\partial H_y/\partial y + \partial H_z/\partial z = 0$ (on peut poser que le champ magnétique est nul suivant l'axe X), cette force se présente sous la forme

$$\bar{f}_z = -\mathfrak{m}_z \frac{\partial H_y}{\partial y}. \tag{36.4}$$

Si l'atome se meut dans le plan ZX avec la vitesse v, par raison de symétrie le vecteur H est contenu dans ce même plan. Il s'ensuit que la force de Lorentz $e/c\,[vH]$ sera dirigée suivant l'axe Y et déterminera un déplacement du pinceau à gauche ou à droite suivant ce même axe. Ces déplacements n'importent nullement ici, la seule chose importante étant le déplacement du pinceau le long de l'axe Z. Mais si la particule s'écarte de Δy du plan ZX, on voit apparaître

aussitôt une composante de la force de Lorentz suivant l'axe Z

$$(f_{\text{Lor}})_z = -\frac{e}{c}\, vH_y.$$

En première approximation $H_y = (\partial H_y/\partial y)\,\Delta y$, de sorte que

$$(f_{\text{Lor}})_z = -\frac{ev}{c}\frac{\partial H_y}{\partial y}\Delta y. \qquad (36.5)$$

Le rapport des déplacements $\Delta_1 z$ et $\Delta_2 z$ de la particule, déterminés par les forces (36.4) et (36.5), s'exprime par

$$\frac{\Delta_1 z}{\Delta_2 z} = \frac{\overline{f_z}}{(f_{\text{Lor}})_z} = \frac{\mathfrak{m}_z}{(ev/c)\,\Delta y}.$$

En assimilant l'atome à un ion simple, on prend pour $\mathfrak{m}_z$ le magnéton de Bohr (35.12) :

$$\frac{\Delta_1 z}{\Delta_2 z} = \frac{\hbar}{2\mu_e v \Delta_y}.$$

En notant μ_a la masse de l'atome, ce rapport s'écrira

$$\frac{\Delta_1 z}{\Delta_2 z} = \frac{1}{4\pi}\frac{\mu_a}{\mu_e}\frac{\lambda}{\Delta y}, \qquad (36.6)$$

où λ est la longueur de l'onde de de Broglie de l'atome : $\lambda = h/\mu_a v$.

En l'absence de champ magnétique, la fente délimitée par les diaphragmes B et B' sera représentée sur la plaque P par une trace horizontale. Lorsqu'on branche un champ magnétique non homogène, le centre de cette trace sera déplacé sous l'action de la force $\overline{f_z}$. Les points non centraux de la trace subissent un déplacement supplémentaire déterminé par l'action de la force de Lorentz. Ce sont les bords de cette trace qui subissent les plus grands déplacements qui se produisent dans des sens contraires. En conséquence, la trace laissée sur la plaque P sera déformée. Pour que la méthode de Stern et Gerlach soit utilisable, il faut que cette déformation de la trace soit petite, ce qui implique que $|\Delta_1 z| \gg |\Delta_2 z|_{\max}$, où $|\Delta_2 z|_{\max}$ dénote le déplacement du bord de la trace provoqué par la force de Lorentz ; c'est donc la valeur correspondant à $|\Delta y| = |\Delta y|_{\max}$. Pour un électron libre $\mu_a = \mu_e$ et on tire de (36.6) $\Delta_1 z/\Delta_2 z = \lambda/(4\pi\,\Delta y)$. Dans ce cas, la condition $|\Delta_1 z| \gg |\Delta_2 z|_{\max}$ ne peut être satisfaite puisqu'on doit avoir $|\Delta y|_{\max} \gg \lambda$. Pour les atomes cette condition peut être satisfaite parce que le rapport μ_a/μ_e est très grand. C'est ce qui est mis à profit dans la méthode de Stern et Gerlach.

§ 37. L'effet Sadovski et le spin du photon

1. En 1889, le physicien russe A. I. Sadovski (1859-1921) a démontré par la théorie qu'une lumière à polarisation circulaire ou elliptique devait posséder un moment cinétique. Pour faire comprendre ce résultat, le plus simple est de se fonder sur la loi de la conservation du moment cinétique. Selon cette loi, le moment cinétique total d'un système fermé doit être constant. Analysons l'effet Sadovski d'abord du *point de vue classique*.

Supposons qu'une charge électrique e tourne sur une circonférence de rayon r autour d'une charge égale et de signe contraire mais

fixe. On sait que pour un mouvement circulaire, l'énergie totale égale à la somme de l'énergie cinétique et de l'énergie potentielle est $\mathcal{E} = -e^2/2r$, donc la moitié de l'énergie potentielle de la charge. Le mouvement circulaire étant un mouvement accéléré, selon l'électrodynamique classique la charge e doit émettre des ondes électromagnétiques. Lorsque la charge rayonne, elle ne peut suivre exactement sa trajectoire circulaire et doit se rapprocher sans cesse de son centre. Supposons qu'au cours d'une seule révolution la diminution de la distance r entre la charge et le centre du cercle est très petite devant le rayon r de ce dernier. On peut alors considérer le mouvement de la charge e comme un mouvement circulaire dont le rayon diminue sans cesse. Les variations concomitantes de l'énergie $\mathcal{E}$ et du rayon r sont liées par la relation

$$d\mathcal{E} = \frac{e^2}{2r^2}\, dr.$$

Le moment cinétique d'une charge tournante est égal à $L = \mu r^2 \omega$, où μ est la masse et ω la pulsation. Pour un mouvement circulaire on a $\mu\omega^2 r = e^2/r^2$, d'où

$$\omega = e/(\mu r^3)^{1/2}.$$

Par conséquent,

$$L = e\,(\mu r)^{1/2},$$

$$dL = \frac{e\mu^{1/2}}{2r^{1/2}}\, dr, \qquad \frac{d\mathcal{E}}{dL} = \frac{e}{(\mu r^3)^{1/2}} = \omega.$$

Ainsi lorsqu'une charge exécute un mouvement circulaire, son énergie et son moment cinétique diminuent de telle manière que leurs variations sont liées par la relation

$$\frac{d\mathcal{E}}{dL} = \omega. \tag{37.1}$$

L'énergie totale et le moment cinétique d'un système fermé doivent rester constants. Le système se compose de substance et de rayonnement, qui peuvent échanger de l'énergie et du moment cinétique. Comme ces grandeurs sont constantes pour le système tout entier, lors d'une variation de r de dr, le rayonnement emporte une énergie $-d\mathcal{E}$ et un moment cinétique $-dL$.

Bien entendu, la structure du rayonnement est déterminée par les processus qui se produisent dans l'émetteur. Mais dès que le rayonnement quitte l'émetteur, tous les liens sont rompus entre eux et le rayonnement existe en tant que système indépendant. Pour ce rayonnement indépendant le rapport entre son énergie et son moment cinétique devient donc une *propriété inhérente au rayonnement lui-même*. Il s'ensuit que dans le cas du procédé d'excitation considéré l'énergie $\mathcal{E}_{ray}$ et le moment cinétique L_{ray} du rayonnement doivent

être liés par la relation

$$\frac{\mathscr{E}_{\text{ray}}}{L_{\text{ray}}} = \omega = \frac{2\pi c}{\lambda}. \qquad (37.2)$$

2. Dans le cas envisagé, la structure du rayonnement émis est assez compliquée. Son intensité et sa polarisation sont différentes suivant les directions. Dans une direction perpendiculaire au plan de la trajectoire circulaire de la charge e, le rayonnement présente une *polarisation circulaire*; dans le plan de ce cercle il est *polarisé rectilignement*, et dans toutes les autres directions sa polarisation est *elliptique*. Bien entendu, on peut transformer la totalité du rayonnement en une onde progressive plane à polarisation circulaire. Pour ce faire, on peut, par exemple, disposer le centre du mouvement circulaire de la charge rayonnante au foyer d'un miroir parabolique infini ayant une réflexion parfaite, le plan du cercle étant normal à l'axe du miroir. On obtient alors une onde réfléchie plane qui progresse parallèlement à l'axe du miroir. Cette onde résulte des interférences des ondes réfléchies de polarisations différentes et présente une polarisation circulaire en raison de la symétrie cylindrique du système. On ne peut cependant affirmer *a priori* que lors de la réflexion dans le miroir le moment cinétique total du rayonnement reste constant.

Pour lever cette difficulté, utilisons un modèle d'émetteur idéalisé, analogue à celui qui fut utilisé dans le tome III, § 83, pour obtenir des ondes électromagnétiques planes. On y avait démontré qu'un plan chargé d'étendue infinie en mouvement accéléré émettait deux ondes électromagnétiques planes se propageant dans des directions différentes avec des énergies égales. Pour éliminer le champ électrostatique des charges, on commençait par placer, à côté du plan considéré, un deuxième plan infini portant des charges de signe opposé, qui était fixe. Modifions ce modèle et considérons à nouveau un plan infini mais fixe et placé dans le vide. Répartissons uniformément sur ce plan des dipôles électriques dont les moments seraient parallèles à ce plan. Supposons que chaque dipôle tourne dans ce plan autour de son centre avec la même vitesse angulaire ω et avec la même phase initiale. Ce plan recouvert de dipôles tournants excitera à nouveau deux ondes planes se propageant dans des directions différentes, mais qui seront polarisées circulairement. Par raison de symétrie l'énergie et le moment cinétique se répartissent également entre les deux ondes. Par conséquent, la relation (37.2) est vérifiée pour chacune de ces ondes prises séparément.

3. De toute évidence le sens de rotation des vecteurs $\boldsymbol{E}$ et $\boldsymbol{H}$ doit coïncider avec le sens de rotation du dipôle $\boldsymbol{p}$ et donc être le même pour les deux ondes. Sur la figure 65 le sens de rotation des dipôles $\boldsymbol{p}$ est indiqué par des flèches. Conformément à ce sens, les moments cinétiques $\boldsymbol{L}_{\text{ray}}$ des deux ondes sont orientés de gauche

à droite. Les deux ondes sont progressives et s'éloignent de la source, ce qui implique que le vecteur de Poynting S des deux ondes est orienté vers l'extérieur. Par conséquent, le vecteur S de l'onde qui s'éloigne vers la droite est orienté vers la droite, et donc dans le sens du vecteur L_{ray}. Le vecteur S de l'onde qui s'éloigne vers la gauche est orienté en sens contraire au sens du vecteur L_{ray}. Mais l'onde qui progresse vers la droite est polarisée circulairement

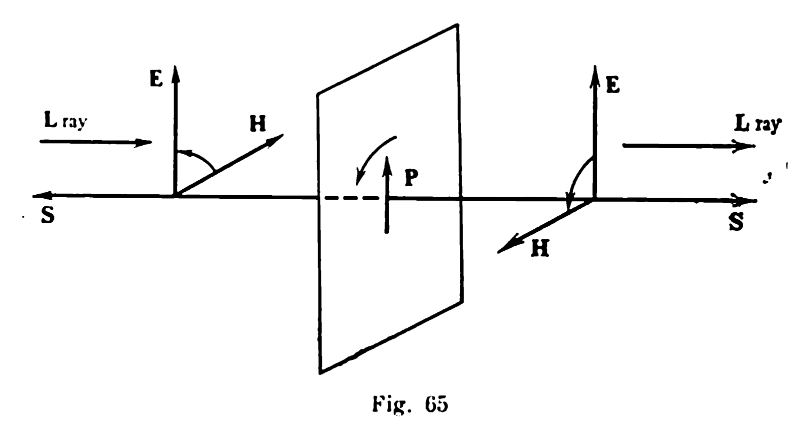


Fig. 65

à gauche (rotation antihoraire des vecteurs E et H si l'onde avance vers l'observateur), tandis que l'onde qui progresse vers la gauche est polarisée circulairement à droite (rotation horaire des vecteurs E et H si l'onde avance vers l'observateur). Par conséquent, dans l'onde polarisée à gauche, le vecteur L_{rav} est orienté dans le sens de propagation de l'onde, et dans l'onde polarisée à droite le vecteur L_{ray} est orienté à l'encontre de son sens de propagation. La même chose vaut pour les ondes à polarisation elliptique.

Il est opportun de remarquer que la disposition réciproque des vecteurs E et H d'une onde progressive détermine univoquement l'orientation du vecteur de Poynting S et donc le sens de propagation de l'onde. Mais cette disposition ne permet pas encore de décider si la polarisation circulaire ou elliptique de l'onde est gauche ou droite. Par exemple, sur les figures 66, *a* et 66, *b*, les vecteurs E, H, S ont la même orientation et les deux ondes se propagent dans le même sens, quoique les sens de rotation des vecteurs E et H (notifiés par des flèches incurvées) y sont opposés: la figure *a* correspond à une polarisation circulaire gauche et la figure *b* à une polarisation circulaire droite.

Voyons maintenant quelle doit être la densité de répartition des dipôles sur le plan émetteur pour que toutes les ondes soient planes et se propagent perpendiculairement au plan émetteur. On doit tenir compte de ce que les ondes émises par les dipôles indi-

viduels interfèrent. Il faut qu'en interférant elles se détruisent mutuellement dans toutes les directions, à l'exclusion de celles qui sont perpendiculaires au plan émetteur. Il suffit pour cela que *la distance entre les dipôles soit inférieure à la longueur d'onde* λ. Dans ce cas, loin du plan émetteur, on verra apparaître des ondes planes qui s'en éloignent. Ce n'est que tout près du plan émetteur

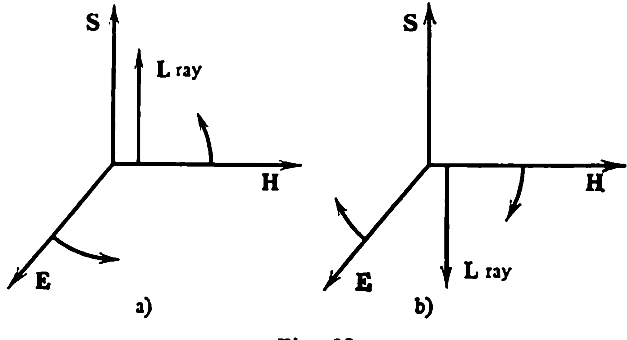


Fig. 66

qu'à ces ondes se superposent des *ondes inhomogènes*, qui ne jouent cependant aucun rôle dans le processus considéré vu qu'elles s'amortissent rapidement dans une couche superficielle mince, d'une épaisseur comparable à la distance de séparation des dipôles.

On arrive ainsi au résultat suivant : *toute onde électromagnétique plane de fréquence* ω, *polarisée circulairement, transporte un moment cinétique lié à l'énergie de l'onde par la relation* (37.2). *Si la polarisation est gauche, le vecteur* L_{ray} *est orienté dans le sens de propagation de l'onde, et si la polarisation est droite, le sens de* L_{ray} *est opposé au sens de propagation.* C'est le principal résultat auquel aboutit A. I. Sadovski.

Le cas de la polarisation elliptique se ramène à celui de la polarisation circulaire, puisque toute onde elliptique peut être décomposée en deux ondes circulaires, l'une polarisée à droite et l'autre à gauche.

4. On peut trouver le moment cinétique du rayonnement par un procédé plus direct en partant des propriétés du *rayonnement lui-même.* On sait que tout rayonnement possède une quantité de mouvement, dont la densité volumique est $g_{el} = S/\omega$, où S est le vecteur de Poynting (cf. t. III, § 81). Si on forme le moment du vecteur g_{el} et si on intègre sur tout l'espace occupé par le rayonnement, on obtient le moment cinétique du rayonnement. Cette procédure est expliquée dans le problème donné à la fin de ce paragraphe, en prenant pour exemple le rayonnement d'un dipôle de Hertz, dont le moment dipôlaire tourne uniformément dans un plan sans changer

de grandeur. Il s'agit bien sûr du moment cinétique de la totalité du rayonnement émis par la source dans toutes les directions. Comme source de rayonnement on peut prendre de nouveau un plan infini recouvert d'un réseau dense de dipôles de Hertz, comme on l'avait fait plus haut. On peut alors obtenir une onde plane polarisée circulairement et introduire pour cette onde la notion de vecteur densité de flux M du moment cinétique du rayonnement. Conformément à (37.2), ce vecteur est défini par la formule

$$M = S/\omega. \tag{37.3}$$

5. Lorsqu'une onde polarisée circulairement est absorbée par un corps, sous incidence normale à l'unité de surface de celui-ci est appliqué un moment de rotation $M = S/\omega$. Pour se faire une idée de la grandeur de l'effet, supposons que la densité du flux d'énergie dans une onde lumineuse plane à polarisation circulaire est $S = 1\ \text{W}\cdot\text{cm}^{-2} = 10^7\ \text{erg}\cdot\text{cm}^{-2}\cdot\text{s}^{-1}$. (C'est une densité près de 7 fois plus grande que celle du rayonnement solaire à la limite de l'atmosphère terrestre.) Pour un rayonnement visible $\lambda = 500$ nm, la formule (37.3) donne

$$M = \frac{S}{\omega} = \frac{\lambda S}{2\pi c} = 2{,}7\cdot 10^{-9}\ \text{dyn/cm},$$

tandis que pour une onde $\lambda = 1$ cm de même intensité lumineuse, on obtient

$$M = 5{,}3\cdot 10^{-5}\ \text{dyn/cm}.$$

Si l'onde traverse une lame cristalline demi-onde, découpée parallèlement à l'axe optique, sa polarisation s'inverse, de droite elle devient gauche et inversement, et à chaque fois la valeur de M double. Pour une puissance de rayonnement donnée l'effet augmente avec la longueur d'onde, mais reste tellement faible qu'il ne fut observé expérimentalement qu'en 1935 par le physicien américain Beth non seulement pour des ondes radio, mais aussi pour la lumière visible.

6. Faisons maintenant l'étude de l'effet Sadovski en nous plaçant au *point de vue quantique*. Ce qui importe ici est que l'émission et la propagation ultérieure de la lumière se font non pas de façon continue mais par *quanta indivisibles*, i.e. par *photons*. En conséquence, il n'est plus nécessaire d'effectuer une concentration artificielle du rayonnement suivant une direction donnée, comme dans le cas de l'étude classique. On n'envisagera pas ici les processus d'émission simultanée de plusieurs photons, car leur probabilité est faible. Une autre particularité de l'approche quantique consiste en ce que *les trois projections sur les axes de coordonnées du vecteur moment cinétique quantique ne peuvent avoir simultanément des valeurs déterminées.*

Lorsque l'atome passe d'un état stationnaire à un autre, ce processus s'accompagne de l'émission d'un *seul photon* d'énergie $\mathcal{E} = \hbar\omega$. La projection du moment cinétique de l'atome sur une direction donnée (axe Z) pour un mouvement orbital de l'électron peut prendre la valeur $m\hbar$. Supposons que lors de l'émission d'un photon cette projection varie de $\hbar$. Cela signifie que pendant l'acte d'émission l'atome perd une énergie $\hbar\omega$ et une projection du moment cinétique égale à $\hbar$. Conformément aux lois de la conservation, l'énergie et le moment cinétique perdus par l'atome sont *emportés par le rayonnement*. On peut en conclure que la projection du moment cinétique du photon émis est égale à $\hbar$. Le moment cinétique interne du photon, i.e. le moment qui n'est pas lié à son mouvement orbital, est appelé *spin du photon*. On dit que le spin du photon est entier et égal à l'unité (en fait il est égal à $\hbar$), quoique la valeur $\hbar$ se rapporte *non pas au moment total* mais seulement *à sa projection sur une direction donnée*. Si la projection (en unités $\hbar$) est égale à s, on peut affirmer, comme généralement pour tout moment cinétique quantique, que le carré du vecteur spin du photon est donné par l'expression $s(s+1)\hbar^2 = 2\hbar^2$. Le rapport des quantités $\mathcal{E} = \hbar\omega$ et $L_z = \hbar$ s'exprime par la formule

$$\frac{\mathcal{E}}{L_z} = \frac{\hbar\omega}{\hbar} = \omega. \tag{37.4}$$

Quoique la forme de cette relation soit la même que celle de la formule classique (37.2), elles sont essentiellement différentes. Dans la formule classique (37.2) L désigne le *moment cinétique total* du rayonnement, tandis que dans la formule quantique (37.4) $L_z = \hbar$ ne représente que la *projection du moment* sur la direction choisie.

7. La masse au repos du photon est égale à zéro. Il ne peut donc exister de référentiel dans lequel le photon serait au repos. Son moment cinétique interne, ou spin, ne peut donc pas être défini comme le moment d'une particule au repos. *Dans tout référentiel, le photon ne peut exister qu'en mouvement avec la vitesse de la lumière c.*

Une étude rigoureuse du moment cinétique du photon ne peut être faite qu'en *théorie quantique relativiste*. Une théorie non relativiste du photon est inconcevable par principe, puisque sa vitesse est toujours égale à la vitesse de la lumière c. Dans ce cours de physique générale, il n'est pas possible d'exposer la théorie relativiste. On se contentera donc de remarquer que, comme toute autre grandeur quantomécanique, le moment cinétique du photon est caractérisé par un *opérateur*. Il a été établi que l'opérateur moment cinétique du photon comporte deux termes. L'un de ces termes est de la forme $[\hat{r}\hat{p}]$, où $\hat{p}$ est l'*opérateur impulsion du photon*, on l'appelle *terme orbital*. Le deuxième terme est l'*opérateur de spin du photon*. La valeur propre de la projection de l'opérateur $[\hat{r}\hat{p}]$ sur la direction

choisie est appelée *moment cinétique orbital du photon*. La valeur propre de la projection de l'opérateur de spin sur la même direction est le *moment cinétique de spin* ou tout court *spin du photon*.

Supposons que le photon ne possède aucun moment orbital et que tout son moment se réduit au spin. Pour fournir une justification évidente de cette assertion, il suffit de remarquer que la longueur de l'onde émise par un atome est généralement très grande devant les dimensions de l'atome. Or le photon ne peut être localisé dans une région de l'espace dont les dimensions linéaires sont inférieures à la longueur d'onde λ de la lumière. D'autre part, comme les dimensions de l'atome sont très petites devant la longueur d'onde λ, l'atome émet le photon pratiquement toujours de façon « centrale » et ce dernier ne recevant aucun moment orbital n'emporte qu'un *moment de spin*. Pour que le photon puisse acquérir encore un moment orbital, le rayonnement doit être émis « obliquement », i.e. par une région périphérique de l'atome se trouvant à une distance de l'ordre de λ du centre. A ces distances, la fonction d'onde de l'atome et la probabilité de l'émission « oblique » du photon sont extrêmement petites.

8. Le fait que le photon n'existe qu'en état de mouvement avec la vitesse c se manifeste aussi en ce que dans tout référentiel il ne peut disposer que d'une *seule direction* — la *direction dans laquelle il se meut*. C'est sur cette direction que se projette le vecteur spin du photon. Mais comme le spin du photon $s = 1$, on pourrait penser que par rapport à cette direction le spin pourrait s'orienter de $2s + 1 = 3$ façons différentes : dans la première la projection du spin est orientée dans le sens du mouvement, dans la deuxième dans le sens opposé et dans la troisième la projection du spin est nulle. En fait la troisième possibilité ne se réalise pas. Notons, sans entrer dans les détails, que l'expérience conduit à cette même conclusion. La transversalité des ondes électromagnétiques implique que pour réaliser n'importe quelle polarisation de l'onde il suffit de superposer deux ondes seulement et non pas trois ondes de polarisations différentes. Conformément au principe de superposition, on doit s'attendre à ce qu'en théorie quantique il soit suffisant de réaliser la superposition de deux états indépendants du photon pour définir n'importe lequel de ses états. Il s'agit donc de savoir quels états du photon peuvent être pris en qualité d'états indépendants. Il suffit pour cela d'étudier la relation qui existe entre la polarisation et le spin du photon.

9. En théorie photonique, la notion de *polarisation* (de même que toutes les autres notions) ne possède pas la signification claire et évidente de la physique classique. Mais comme la polarisation de la lumière existe en tant que phénomène observable, il importe de mettre en évidence ce qui lui correspond dans la théorie photonique. *La seule grandeur orientée caractérisant les propriétés internes du*

photon est le spin. D'autre part, en théorie classique, le moment cinétique L d'une onde polarisée circulairement est orienté dans le sens de propagation de l'onde ou à l'encontre de celle-ci. *Il est donc tout indiqué de considérer que le photon est polarisé circulairement lorsqu'il se trouve dans un état caractérisé par une valeur déterminée de la projection du spin sur sa direction de propagation.* Si le spin est orienté dans le sens de propagation de la lumière, on dira que le photon est polarisé *à gauche*, et dans le cas contraire il présentera une polarisation *droite* *).

En optique classique, on peut réaliser n'importe quelle polarisation (rectiligne ou elliptique) d'une onde progressive plane par superposition de deux ondes planes (cohérentes) se propageant dans la même direction ayant l'une une polarisation gauche et l'autre une polarisation droite. De même *l'état d'un photon polarisé circulairement et se propageant dans une direction déterminée doit être considéré comme un état propre auquel correspondent les valeurs propres de la projection du spin, égales à* $s_z = +1, 0, -1$. *Par superposition linéaire de ces états, on peut obtenir une polarisation quelconque du photon.* Mais l'état avec $s_z = 0$ ne se réalise pas.

C'est pourquoi, *l'état du photon ayant une polarisation quelconque et se propageant dans une direction déterminée peut être réalisé par superposition linéaire de deux états seulement: l'état avec* $s_z = +1$ *et l'état avec* $s_z = -1$. Bien entendu, la superposition de tels états n'est pas une superposition *classique*; on doit l'entendre dans le même sens que la superposition d'états quantiques d'une particule caractérisés par des fonctions d'ondes. Comme les états d'un photon avec $s_z = +1$ et $s_z = -1$ sont des états *propres*, les carrés des modules des coefficients figurant devant ces états dans une superposition déterminent les *probabilités relatives* des états eux-mêmes. Cela se manifeste, par exemple, dans le fait que la mesure de la projection s_z (par mesure du moment rotatoire qu'acquiert un corps à la suite de l'absorption d'un photon) peut fournir, avec des probabilités adéquates, soit la valeur $s_z = +1$, soit la valeur $s_z = -1$. Aucun autre résultat ne peut être obtenu.

PROBLÈME

Un dipôle rigide de moment électrique p est animé d'un mouvement de rotation uniforme autour de son centre, avec la vitesse angulaire constante ω, le vecteur p étant toujours contenu dans un même plan. Selon l'électrodynamique classique, ce dipôle rayonne comme un dipôle de Hertz. Le rayonnement émis possède une quantité de mouvement dont la densité volumique est donnée

*) Cette définition des polarisations gauche et droite correspond à la convention en vigueur en optique classique. En électrodynamique quantique, on a adopté la convention contraire : à la polarisation droite correspond l'orientation du spin coïncidant avec le sens de propagation du photon et à la polarisation gauche correspond l'orientation opposée du spin.

par l'expression (cf. t. III, § 84)

$$g_{el} = \frac{1}{4\pi c} [EH]. \qquad (37.5)$$

Le rayonnement emporte avec lui un moment cinétique. Pour calculer le moment total emporté par le rayonnement, il suffit de connaître g_{el} sur une sphère infiniment éloignée dont le centre est confondu avec le centre du dipôle. Discuter la question suivante: de quelle façon peut apparaître le moment, sachant que loin du dipôle (dans la zone d'onde) les champs E et H sont liés l'un à l'autre comme dans une onde plane et sont tous deux perpendiculaires au rayon vecteur r mené du dipôle au point d'observation? Dans ces conditions, le vecteur g_{el} est dirigé suivant r et par suite le moment $[rg_{el}]$ est nul.

Solution. L'assertion que dans la zone d'onde les champs E et H décroissent avec la distance comme $1/r$ est approchée et n'est vérifiée qu'asymptotiquement lorsque $r \to \infty$. Cette approximation est suffisante lorsqu'il s'agit de calculer l'énergie totale ou l'impulsion totale emportées par le rayonnement, puisque dans ces cas le vecteur $[EH]$ varie comme $1/r^2$. Il est inutile de tenir compte des puissances supérieures de la quantité $1/r$ puisque, lors de l'intégration sur une sphère, elles n'apportent aucune contribution lorsqu'on passe à la limite $r \to \infty$. Par contre, la densité du moment cinétique $l_{el} = [rg_{el}]$ se déduit de g_{el} en le multipliant vectoriellement par r. Il suffit, bien entendu, de calculer g_{el} sur la sphère éloignée aux termes $1/r^2$ près et pour cela il faut tenir compte, dans l'expression de g_{el}, des termes de troisième degré en $(1/r)$. A cet effet, on utilisera les formules du champ de rayonnement du dipôle de Hertz dans la zone d'onde dans le vide:

$$E = \left[\frac{3(\dot{p}r)}{cr^4} r - \frac{\dot{p}}{cr^2}\right]_{t-r/c} - \left[\frac{(\ddot{p}r)}{c^2r^3} r - \frac{\ddot{p}}{c^2r}\right]_{t-r/c},$$
$$H = \frac{1}{cr^3}[\dot{p}r]_{t-r/c} + \frac{1}{c^2r^2}[\ddot{p}r]_{t-r/c}. \qquad (37.6)$$

Ces formules se déduisent des formules (141.10) du tome III écrites pour le cas du vide, ce qui conduit à poser $D = E$ et $v = c$. Dans la première des formules (141.10) on a négligé le premier terme proportionnel à $1/r^3$, puisqu'il ne peut influencer la valeur de g_{el} qu'en affectant un terme qui est de l'ordre de $1/r^4$. A l'aide des formules (37.6) on doit déduire le produit $[EH]$ dans l'approximation requise, tout en négligeant les termes colinéaires avec r, vu qu'ils ne contribuent pas au calcul de $[rg_{el}]$. En omettant l'indice $t - r/c$ on obtient ainsi

$$g_{el} = \frac{1}{2\pi c^4 r^4} (\dot{p}r)\ddot{p} + (\ldots) r, \qquad (37.7)$$

$$l_{el} = \frac{1}{2\pi c^4 r^4} (\dot{p}r)[r\ddot{p}]. \qquad (37.8)$$

On peut transformer cette dernière formule en mettant à profit le fait que le vecteur p ne modifie pas sa longueur et ne varie que par suite de la rotation, ce qui implique que $\dot{p} = [\omega p]$. La même remarque s'applique à $\dot{p}$, i.e. $\ddot{p} = [\omega \dot{p}]$. La formule (37.8) se transforme ainsi en la formule

$$l_{el} = \frac{1}{2\pi c^4 r^4} (\dot{p}r)[r[\omega\dot{p}]] = \frac{1}{2\pi c^4 r^4} (\dot{p}r)\{(\dot{p}r)\omega - (\omega r)\dot{p}\}. \qquad (37.9)$$

Pour trouver le moment de rayonnement total émis par le dipôle dans l'unité de temps, il faut multiplier cette expression par c et intégrer le résultat sur toute la surface d'une sphère infiniment lointaine. En raison de la symétrie de rotation autour de ω, cette intégration fournit, tout naturellement, un vecteur orienté

suivant ω. Comme le vecteur $\dot{p} = [\omega p]$ est orthogonal à ω, on peut omettre le dernier terme de (37.9). On obtient alors

$$l_{el} = \frac{1}{2\pi c^4 r^4} (\dot{p} r)^2 \omega. \qquad (37.10)$$

On peut intégrer comme si le vecteur $\dot{p}$ était fixe et adopter le système de coordonnées sphérique représenté sur la figure 67. Dans ce cas, l'élément de surface de la sphère sera égal à $dS = r^2 \sin\vartheta \, d\vartheta \, d\varphi$. Finalement, le moment d'impulsion de la quantité du rayonnement est

$$L_{ray} = \frac{\omega \dot{p}}{2\pi c^3} \int_0^{\pi} \int_0^{2\pi} \sin^3\vartheta \, d\vartheta \, d\varphi = \frac{2}{3c^3} \dot{p}^2 \omega. \qquad (37.11)$$

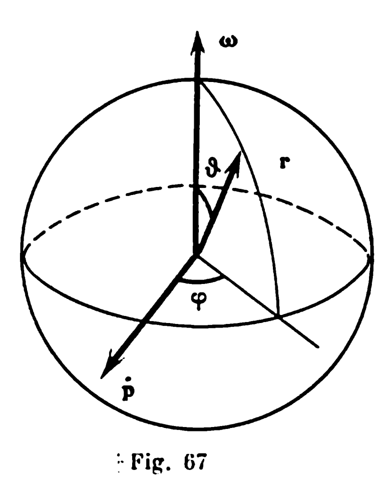


Fig. 67

L'énergie rayonnée par le dipôle dans l'unité de temps est $\mathscr{E}_{ray} = \frac{2}{3c^3} \ddot{p}^2$ (cf. t. III, § 141). Comme $\ddot{p}^2 = \omega^2 \dot{p}^2$, on obtient

$$L_{ray} = \mathscr{E}_{ray}/\omega. \qquad (37.12)$$

§ 38. Les quatre nombres quantiques de l'électron et la structure fine des termes spectraux

1. Le fait que l'électron possède un moment cinétique propre (le spin) signifie qu'on ne peut caractériser l'état de l'électron (à la différence d'une particule ponctuelle classique) à l'aide de trois degrés de liberté seulement. L'électron dispose dans l'atome d'un quatrième degré de liberté appelé *degré de liberté de spin*. On notera qu'on n'envisage provisoirement que l'atome hydrogénoïde, ou plus généralement un atome ou un ion à plusieurs électrons, mais avec un *seul* électron périphérique (électron de valence ou optique). C'est cet électron périphérique qui fera l'objet de notre étude. En mécanique quantique, l'état de cet électron est caractérisé par quatre nombres quantiques: 1) le *nombre principal n*, 2) le *nombre orbital l*, 3) le *nombre magnétique orbital* m_l, 4) le *nombre de spin* m_s. On a déjà précisé la signification des trois premiers nombres quantiques n, l et m_l au § 33. Le nombre de spin m_s détermine les projections du vecteur spin s sur la direction choisie. Si l'atome se trouve déjà dans un état caractérisé par une valeur déterminée du moment orbital l (i.e. avec l^2 et l_z déterminés), la direction privilégiée (axe Z) est déterminée par le vecteur l si $l^2 \neq 0$. Le spin s peut être orienté

parallèlement ou antiparallèlement à $\boldsymbol{l}$. Cela signifie que la projection du vecteur $\boldsymbol{s}$ sur la direction choisie ne peut prendre que deux valeurs: $+\hbar/2$ et $-\hbar/2$, ou ce qui revient au même, les valeurs $m_s\hbar$ avec $m_s = \pm{}^1/_2$. Lorsque $l = 0$ (l'atome se trouve alors dans l'état s), tout le moment cinétique de l'atome est dû uniquement au spin : $\boldsymbol{s}$. Si l'atome se trouve dans un état tel que l'une des projections s_x, s_y, s_z a une valeur déterminée (égale à $\pm\hbar/2$), l'axe correspondant est la direction privilégiée dans l'atome.

2. Le moment cinétique orbital $\boldsymbol{l}$ et le moment de spin $\boldsymbol{s}$ se composent vectoriellement pour former le moment cinétique total $\boldsymbol{j} = \boldsymbol{l} + \boldsymbol{s}$ (cf. § 32). La projection du moment total sur la direction choisie peut prendre les valeurs $m_j\hbar$, où $m_j = m_l + m_s = m_l \pm {}^1/_2$ s'appelle *nombre quantique de la projection du moment total.* Il est clair que les opérateurs des projections du moment total sur les axes de coordonnées vérifient les mêmes relations de commutation (31.6) que les opérateurs des projections du moment orbital. Il s'ensuit que dans un état donné seuls le carré du moment total $\boldsymbol{j}^2$ et l'une de ses projections sur les axes de coordonnées peuvent avoir des valeurs déterminées. Une autre conséquence est que $\boldsymbol{j}^2 = \hbar^2 j\,(j+1)$, où j est la valeur maximale que peut prendre le nombre quantique m_j. Parfois on désigne j sous le nom de *nombre quantique interne.*

Puisque j est la valeur maximale du nombre m_j, et l est la valeur maximale du nombre m_l, il résulte de la relation $m_j = m_l \pm {}^1/_2$ que

$$j = l \pm {}^1/_2. \tag{38.1}$$

Le signe « plus » correspond au cas où le spin de l'électron est parallèle au moment orbital, et le signe « moins » au cas où il lui est antiparallèle. Dans les deux cas j est un nombre demi-entier, puisque l est toujours entier.

Pour une valeur donnée de j, il peut y avoir $2j + 1$ états quantiques différant les uns des autres par les valeurs du nombre quantique m_j:

$$m_j = -j,\ -(j-1),\ \ldots,\ +(j-1),\ +j. \tag{38.2}$$

Dans le cas d'atomes n'ayant qu'un seul électron de valence, le nombre de ces états est toujours pair puisque j est demi-entier.

3. Pour caractériser les états d'un atome monoélectronique, on peut remplacer les nombres quantiques n, l, m_l, m_s par d'autres groupes de quatre nombres quantiques, par exemple n, l, j, m_s. En spectroscopie, il est d'usage d'utiliser les nombres quantiques n, l, j, $2s + 1$ et de remplacer le nombre l par la lettre correspondante de l'alphabet latin, conformément au tableau du § 34, 2. On écrit d'abord la valeur numérique du nombre quantique principal n, puis la lettre qui remplace le nombre l, en écrivant le nombre j en indice du bas à droite de cette lettre, et le nombre $2s + 1$ en

indice du haut à gauche de cette lettre. Le nombre $2s + 1$, appelé *multiplicité du niveau*, sert à indiquer le nombre d'orientations différentes du spin par rapport à l'orientation du moment orbital l. Pour un atome n'ayant qu'un seul électron de valence, l'indice $2s + 1$ peut être supprimé puisque pour l'électron $s = 1/2$, et par suite on a toujours $2s + 1 = 2$. Mais si le spin de la particule avait une valeur différente, l'indication de la multiplicité $2s + 1$ revêtirait une grande importance. Il en va de même pour le nombre correspondant pour les atomes à plusieurs électrons de valence (voir sous-paragraphe 10).

Considérons, par exemple, l'état

$$3^2s_{1/2}$$

(ce qui se lit: « trois, doublet $s_{1/2}$ ». La signification de ce mode de lecture est explicitée en ci-dessous). Dans cet état, $n = 3$, $l = 0$, $j = 1/2$. Le moment cinétique total n'est dû qu'au spin. Dans la formule $j = l \pm 1/2$, le signe « moins » doit être rejeté puisque $j > 0$, et, par suite, $j = l + 1/2$. Cet état est appelé doublet de façon purement formelle, car pour $l = 0$ toutes les directions sont équivalentes pour l'orientation du spin. En fait il s'agit d'un singulet. Ces considérations concernent naturellement tous les états s (i.e. les états avec $l = 0$).

Prenons pour un autre exemple l'état dit « quatre, doublet $d_{3/2}$ », i.e.

$$4\,^2d_{3/2}.$$

Dans cet état, $n = 4$, $l = 2$, $j = 3/2$, avec $j = l - 1/2$, ce qui signifie que le moment de spin est antiparallèle au moment orbital. Par contre, dans l'état $4\,^2d_{5/2}$, $j = l + 1/2$, et le spin est parallèle au moment orbital. Ainsi, l'état d est un doublet véritable. La même chose vaut pour tous les autres états: p, f, g, . . . (à l'exclusion du seul état s).

4. La principale interaction entre l'électron et le noyau est l'*interaction électrostatique* de leurs charges. Mais comme l'électron est en mouvement par rapport au noyau, il apparaît une interaction supplémentaire déterminée par le spin de l'électron et la charge du noyau, appelée *interaction spin-orbite*. Cette interaction apparaît de façon évidente en faisant appel aux conceptions de la théorie semi-classique de Bohr. Le modèle le plus simple qui soit est le modèle de l'atome d'hydrogène où l'électron se meut sur une orbite circulaire. Rapportons ce modèle à un référentiel dans lequel l'électron est au repos, ce qui signifie que le référentiel se meut en même temps que l'électron. Dans ce référentiel, le noyau est en mouvement et crée donc un champ magnétique H qui exerce son action sur le moment magnétique de spin $\vec{\mathfrak{m}}_s$ de l'électron qui est au repos dans ce référentiel. Comme les charges du proton et de l'électron sont

égales numériquement et de signes contraires, le noyau en mouvement dans le référentiel mobile produit à l'emplacement de l'électron le même champ magnétique que celui qui serait créé par l'électron en rotation dans un référentiel fixe à l'emplacement du noyau. On peut donc assimiler formellement l'interaction spin-orbite à une interaction entre les moments magnétiques de spin et orbital de l'électron.

Le moment magnétique de spin $\vec{\mathfrak{m}}_s$ de l'électron peut s'orienter parallèlement ou antiparallèlement au champ magnétique orbital. Dans le premier cas, l'énergie potentielle d'interaction de l'électron avec le noyau augmente, et dans le second cas elle diminue. Par suite de l'interaction spin-orbite, chaque niveau énergétique de l'atome se scinde en deux sous-niveaux. Cela ne se produit pas lorsque l'atome se trouve dans un état s, puisque alors il ne possède pas de moment magnétique orbital et il ne peut y avoir d'interaction spin-orbite. L'éclatement des niveaux d'énergie résultant de l'interaction spin-orbite porte le nom de *structure fine des niveaux*. L'ensemble des sous-niveaux formés à partir d'un niveau donné constitue un *multiplet*. Suivant le nombre de sous-niveaux du multiplet, on distingue les *doublets*, les *triplets*, les *quadruplets*, les *quintuplets*, etc. Les niveaux simples qui n'éclatent pas sont appelés des *singulets*. Les mêmes termes servent à désigner les ensembles de raies spectrales résultant de l'éclatement d'une raie (cf. § 40).

Ainsi, dans le cas d'atomes ou d'ions ne possédant qu'un seul électron de valence, l'interaction spin-orbite provoque l'éclatement en doublets de tous les niveaux énergétiques, à l'exclusion du niveau s qui reste toujours singulet. Le sens des notations introduites cidessus au sous-paragraphe 3 est maintenant parfaitement clair. Par exemple, le niveau $4\,{}^2d_{3/2}$ fut énoncé « quatre, doublet $d_{3/2}$ ». On a déjà signalé que l'utilisation du terme doublet pour désigner les niveaux singulets s est toute conventionnelle et n'est utilisée que pour ne pas singulariser ces niveaux parmi les niveaux doublets véritables p, d, f, . . . D'ailleurs, on peut considérer de façon formelle les niveaux singulets s comme des doublets formés par deux sous-niveaux confondus. La signification physique de la multiplicité $2s + 1$ est également claire : elle détermine le nombre de sous-niveaux dans les multiplets apparaissant à la suite des interactions spin-orbite.

5. On peut estimer sans difficulté l'ordre de grandeur de l'énergie potentielle supplémentaire due à l'interaction spin-orbite. Prenons un atome d'hydrogène à l'état fondamental et appliquons le mécanisme d'apparition de cette interaction, qui a été décrit plus haut. Reprenons le référentiel qui se meut avec l'électron. Le champ magnétique créé à l'emplacement de l'électron par le proton animé d'une vitesse v est donné par la formule $H = \frac{e[vr]}{cr^3}$, où r est le rayon vec-

teur de l'électron par rapport au proton. En valeur absolue $H = \alpha e/r^2$ avec $\alpha = v/c$. Selon (13.19) cette dernière quantité est la *constante de la structure fine* définie par (13.8), i.e. $\alpha = e^2/\hbar c$. Dans le champ magnétique H, l'électron possède l'énergie potentielle $-(\vec{\mathfrak{m}}_s H)$, le vecteur $\vec{\mathfrak{m}}_s$ pouvant être parallèle ou antiparallèle à H. En valeur absolue, cette énergie est égale à $\mathfrak{m}_s H = \mathfrak{m}_B H$, où $\mathfrak{m}_B = e\hbar/(2\mu_e c)$ est le magnéton de Bohr. Comparons-la à l'énergie totale de l'atome d'hydrogène dans l'état fondamental. Selon la formule (13.20), cette énergie est donnée par l'expression $\mathscr{E}_1 = -{}^1/_2\alpha^2\mu_e c^2$. On doit poser r égal au rayon de Bohr défini par la formule (13.16), i.e. $r_B = \hbar^2/(\mu_e c^2)$. On obtient finalement

$$\frac{\mathfrak{m}_B H}{\mathscr{E}_1} = \alpha^2 = 5{,}325 \cdot 10^{-5}. \tag{38.3}$$

6. Puisque $\alpha = v/c$ (v est la vitesse que possède l'électron sur la première orbite de Bohr), *l'interaction spin-orbite est un effet quadratique par rapport au paramètre* α. Par conséquent, la théorie de cet effet ne peut être que *relativiste*. On pouvait s'y attendre puisque *le spin lui-même est un effet quantique relativiste, qui disparaît dans l'approximation non relativiste. La variation de la masse avec la vitesse conduit aussi à l'éclatement fin des niveaux d'énergie* même dans le cadre de la théorie semi-classique de Bohr, comme le démontra Sommerfeld. La raison en est que dans la théorie non relativiste de Bohr, à toutes les orbites elliptiques de l'électron (l'orbite circulaire y comprise) ayant même grand axe correspond une *seule et même énergie*. Dès qu'on tient compte de la variation de la masse avec la vitesse, *cette dégénérescence est levée*, et la valeur de l'énergie devient fonction de l'excentricité de l'ellipse. C'est ce qui détermine l'éclatement fin du niveau énergétique. On peut donc donner une définition plus exacte de la structure fine en disant qu'*elle est déterminée aussi bien par l'interaction spin-orbite que par la dépendance de la masse de l'électron avec la vitesse*. Comme les deux causes d'éclatement sont du deuxième ordre par rapport au paramètre α, on doit en tenir compte simultanément.

La méthode de calcul et d'étude la plus conséquente de la structure fine est fondée sur la *théorie quantique relativiste de Dirac*, qui tient automatiquement compte du spin de l'électron et de la variation de sa masse avec la vitesse.

Pour un atome hydrogénoïde, la solution de l'équation d'onde relativiste de Dirac conduit à la formule suivante de l'énergie de l'état stationnaire:

$$\mathscr{E} = -\frac{(Ze^2)^2\mu_e}{2\hbar^2 n^2}\left[1 + \frac{\alpha^2 Z^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4n}\right)\right]. \tag{38.4}$$

Dans l'expression entre crochets on a omis les termes de la quatrième et des puissances supérieures par rapport à α. Par suite de la

petitesse de la constante α^2 la correction à apporter à la formule non relativiste (13.8) est très petite, de sorte que l'éclatement considéré des niveaux est appelé à juste titre « structure fine ».

Notons que selon la théorie de Dirac, les énergies des niveaux des atomes hydrogénoïdes sont dégénérées en l, autrement dit que ces énergies ne dépendent (à n'importe quelle approximation) que du nombre quantique principal n et du nombre quantique du moment total j, tout en étant indépendantes du nombre quantique orbital l (les écarts de ce résultat sont considérés au § 44). Cela signifie que pour l'atome d'hydrogène et les atomes hydrogénoïdes les niveaux caractérisés par les mêmes nombres quantiques n et j et des l différents coïncident. Cela ne se produit que pour l'hydrogène et les atomes hydrogénoïdes. Pour tous les autres atomes monoélectroniques, les atomes des métaux alcalins par exemple, les niveaux ne coïncident pas.

7. Dans les atomes légers la valeur de l'éclatement fin des niveaux énergétiques ne dépasse pas 10^{-5} eV, mais croît notablement avec la charge du noyau. Pour les atomes lourds, la valeur de l'éclatement peut atteindre quelques dixièmes d'électron-volt et il ne peut donc être question dans ce cas de parler de structure fine. (Rappelons que l'énergie d'ionisation de l'atome d'hydrogène pris à l'état fondamental est égale à 13,6 eV.)

Précisons encore qu'outre la structure fine on décèle encore dans les spectres de l'hydrogène et de nombreux autres atomes une *structure* dite *hyperfine*. Cette dernière résulte de l'interaction des moments magnétiques des électrons avec les champs magnétiques faibles créés par les noyaux atomiques. Aucune formule analogue à (38.4) ne permet de décrire les structures hyperfines. La structure hyperfine sera considérée en Physique nucléaire.

8. Afin de n'avoir pas à revenir sur la question des nombres quantiques et de ne pas avoir à exposer deux fois les règles de sélection relatives à l'émission de la lumière, qui font l'objet du paragraphe suivant, il convient d'envisager succinctement la question des *atomes à plusieurs électrons*. Une étude détaillée de ces questions que nous ne ferons qu'effleurer est donnée dans des cours de spectroscopie spéciaux. Dans un cours de physique générale, on ne peut en donner que quelques notions préliminaires, approchées d'ailleurs.

Dans le cas d'atomes à plusieurs électrons, chaque (i-ième) électron de l'enveloppe électronique de l'atome peut être caractérisé par les vecteurs moments cinétiques orbital $\boldsymbol{l}_i$ et de spin $\boldsymbol{s}_i$. Or l'expérience montre que dans l'étude des questions les plus importantes, on peut se passer d'une caractéristique aussi complète et combiner d'une certaine façon, en appliquant la règle d'addition vectorielle, les moments orbitaux et de spin des différents électrons. Si on n'avait besoin de connaître que le moment cinétique total $\boldsymbol{J}$ de l'atome, l'ordre de composition des vecteurs $\boldsymbol{l}_i$ et $\boldsymbol{s}_i$ importerait

peu, puisque le résultat final est alors indépendant de l'ordre des termes. En fait, on doit connaître, en plus de $\boldsymbol{J}$, plusieurs autres moments et les nombres quantiques correspondants. Ces moments dérivent de $\boldsymbol{l}_i$ et $\boldsymbol{s}_i$ en sélectionnant certains groupes de termes. Quels groupes de termes faut-il sélectionner pour y additionner les $\boldsymbol{l}_i$ et $\boldsymbol{s}_i$? Cela dépend des intensités relatives des différentes interactions entre les électrons de l'atome. La plus importante est la *liaison normale* ou *liaison Russel-Saunders*, astrophysiciens américains qui la définirent en 1925. Cette liaison s'établit lorsque l'interaction électrostatique des électrons (leur répulsion coulombienne) est grande par rapport à l'interaction spin-orbite (interaction entre les moments magnétiques orbitaux et de spin des électrons). C'est ce qui se réalise généralement dans les atomes légers et de masses moyennes.

9. La liaison normale consiste en ce que les moments orbitaux et de spin des électrons de l'enveloppe électronique s'additionnent séparément suivant les règles de composition vectorielle pour donner les *moments résultants orbital* et *de spin de l'atome* que l'on dénote par les lettres majuscules $\boldsymbol{L}$ et $\boldsymbol{S}$, i.e.

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{l}_1 + \boldsymbol{l}_2 + \boldsymbol{l}_3 + \ldots, \\ \boldsymbol{S} &= \boldsymbol{s}_1 + \boldsymbol{s}_2 + \boldsymbol{s}_3 + \ldots \end{aligned} \qquad (38.5)$$

On caractérise l'état de l'enveloppe électronique d'un atome par les moments globaux $\boldsymbol{L}$ et $\boldsymbol{S}$, ainsi que par le moment cinétique total de l'atome qui, de toute évidence, dépend de l'angle que font entre eux les vecteurs $\boldsymbol{L}$ et $\boldsymbol{S}$. Le moment cinétique total est donné par la formule

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}. \qquad (38.6)$$

Aux vecteurs $\boldsymbol{J}$, $\boldsymbol{L}$, $\boldsymbol{S}$ correspondent les nombres quantiques J, L, S qui déterminent les carrés des modules de ces vecteurs (en unités $\hbar$) selon les formules

$$\begin{aligned} \boldsymbol{J}^2 &= J\,(J + 1), \\ \boldsymbol{L}^2 &= L\,(L + 1), \\ \boldsymbol{S}^2 &= S\,(S + 1). \end{aligned} \qquad (38.7)$$

Il est bien évident que lorsque l'atome comporte un nombre d'électrons pair, les nombres quantiques S et J sont entiers et lorsque ce nombre est impair, les nombres S et J sont demi-entiers. Le nombre quantique L est toujours entier. Comme toujours, les nombres J, L, S représentent les plus grandes valeurs que peuvent avoir les projections des vecteurs $\boldsymbol{J}$, $\boldsymbol{L}$, $\boldsymbol{S}$ sur une direction prédéterminée. Il s'ensuit que les projections de ces vecteurs peuvent prendre les

valeurs suivantes (en unités $\hbar$):

$$\begin{aligned} m_J &= -J,\ -(J-1),\ \ldots,\ +(J-1),\ +J, \\ m_L &= -L,\ -(L-1),\ \ldots,\ +(L-1),\ +L, \\ m_S &= -S,\ -(S-1),\ \ldots,\ +(S-1),\ +S. \end{aligned} \quad (38.8)$$

Lorsque L et S sont donnés, le nombre quantique J peut prendre les valeurs suivantes:

$$J = |L+S|,\ |L+S-1|,\ \ldots,\ |L-S|. \quad (38.9)$$

Il est clair que pour calculer les vecteurs L, S, J il suffit de prendre en considération les seuls électrons de valence, à condition toutefois que les couches électroniques internes de l'atome *soient entièrement remplies*, puisque alors les moments cinétiques orbitaux et de spin des électrons des couches internes *se compensent* mutuellement, i.e. leurs moments totaux sont nuls.

Comme les électrons des atomes sont soumis à l'action du champ électrique du noyau ayant une symétrie centrale, le vecteur moment total J se conserve intégralement. Quant aux vecteurs L et S ils ne se conservent pas séparément et varient par suite des interactions spin-orbite, quoique leurs longueurs et, par suite, les nombres quantiques L et S restent pratiquement inchangés. Les projections des vecteurs L et S sur la direction du vecteur J se conservent pratiquement elles aussi. De ce fait on peut se représenter la variation temporelle de L et S comme une précession (rotation) de ces vecteurs autour de la direction invariable du vecteur J, et ce avec la même vitesse angulaire. Ce comportement présente une certaine analogie avec la précession libre de l'axe de la figure et de la vitesse angulaire ω d'un gyroscope symétrique autour de la direction invariable du vecteur moment cinétique (cf. t. I, § 49). La différence entre les deux systèmes réside en ce que dans le cas du gyroscope les directions de l'axe de la figure et du vecteur ω peuvent varier de façon continue, tandis qu'elles sont quantifiées dans l'atome. Cela tient à ce que les projections des vecteurs L et S sur la direction du vecteur J ne peuvent prendre que les valeurs quantifiées $m_L\hbar$ et $m_S\hbar$, où m_L et m_S sont les nombres quantiques correspondants ne pouvant prendre que les valeurs données par les formules (37.8).

10. En spectroscopie l'état des électrons de valence de l'atome est caractérisé sommairement par le nombre quantique L, avec la particularité qu'on substitue à la valeur numérique de L une des lettres de l'alphabet latin, exactement comme on le fait dans le cas d'un seul électron de valence (cf. § 34, 2), mais en utilisant des lettres *majuscules*. Autrement dit, on utilise les notations indiquées dans le tableau suivant:

Nombre quantique L	0	1	2	3	4	5	6	7	8
Notation de l'état	S	P	D	F	G	H	I	K	L

et les lettres suivantes en omettant les lettres P et S.

On affecte ces lettres majuscules d'un indice placé à droite et en bas et représentant le nombre quantique du moment total J et d'un second indice placé à gauche et en haut et représentant le nombre $2S + 1$ appelé *multiplicité du niveau*. Connaissant ce nombre on peut calculer aussi bien le spin S que le nombre de sous-niveaux formés par éclatement du niveau considéré par suite de l'interaction spin-orbite. On notera que le nombre $2S + 1$ n'indique le nombre de composantes formées par éclatement du niveau considéré que lorsque $S \leqslant L$. Si $S \geqslant L$, le nombre de composantes formées par éclatement est déterminé par le nombre de projections possibles du vecteur L sur le vecteur S dont la longueur est plus grande, et sera donc égal à $2L + 1$. Mais même dans ce cas, quoique formellement, le nombre $2S + 1$ est appelé multiplicité du niveau.

Par exemple, lorsque la couche électronique externe de l'atome contient deux électrons, deux cas sont à envisager: 1) les spins des électrons sont antiparallèles et, par suite, $S = 0$; 2) les spins des électrons sont parallèles et $S = 1$.

Dans le premier cas $J = L$, $2S + 1 = 1$, ce qui signifie que tous les niveaux sont des *singulets*. Suivant les valeurs de L on obtient les niveaux suivants:

$L=J$	0	1	2	3	4	5	6
Niveaux	$^1S_0$	$^1P_1$	$^1D_2$	$^1F_3$	$^1G_4$	$^1H_5$	$^1I_6$

Dans le second cas $2S + 1 = 3$ et tous les niveaux sont des *triplets*, à l'exclusion des niveaux s qui sont toujours des *singulets*. Trois possibilités existent alors: $J = L - 1$, $J = L$, $J = L + 1$, ce qui conduit aux résultats consignés dans le tableau suivant:

L	0		1			2			3			4		
J	0	1	0	1	2	1	2	3	2	3	4	3	4	:
Niveaux	$^3S_0$	$^3S_1$	$^3P_0$	$^3P_1$	$^3P_2$	$^3D_1$	$^3D_2$	$^3D_3$	$^3F_2$	$^3F_3$	$^3F_4$	$^3G_3$	$^3G_4$	$^3G_5$

Le lecteur est invité à considérer le cas où la couche externe de l'atome contient trois électrons.

C'est bien évident qu'on n'arrive pas à caractériser pleinement l'état de l'enveloppe électronique d'un atome en indiquant les nombres quantiques J, L, S. En spectroscopie on indique encore la configuration électronique de la couche externe de l'atome, i.e. les nombres d'électrons se trouvant dans les états s, p, d, . . .

11. Notons encore une fois que *la liaison dite normale n'est pas la seule possible. car ce n'est qu'un cas limite.* L'autre cas limite est la liaison dite (j, j) qui se manifeste lorsque l'interaction spin-orbite est grande par rapport aux interactions électrostatiques mutuelles des différents électrons. Dans le cas de cette liaison. les moments orbital et de spin de chaque électron s'additionnent en donnant le moment total $\boldsymbol{j}_i = \boldsymbol{l}_i + \boldsymbol{s}_i$. On caractérise l'état de l'enveloppe électronique de l'atome par ces moments et par les nombres quantiques correspondants. Il est clair que le moment total $\boldsymbol{J}$ de l'atome tout entier ne dépend pas de l'ordre de disposition des termes $\boldsymbol{l}_i$ et $\boldsymbol{s}_i$ et peut être calculé en prenant leur somme vectorielle par la formule

$$\boldsymbol{J} = \sum \boldsymbol{j}_i. \tag{38.10}$$

Une liaison (j, j) nette se rencontre dans les atomes lourds, quoique assez rarement. Différents autres types intermédiaires de liaisons se manifestent dans les atomes. Dans ce cours on n'aura affaire qu'à la liaison normale qui est la plus importante.

§ 39. Règles de sélection pour l'émission et l'absorption de la lumière

1. Lorsqu'un atome se trouve dans un état stationnaire excité, il peut transiter dans un état de plus faible énergie avec émission d'un photon. Inversement, un atome peut absorber un photon et passer sur un niveau de plus grande énergie. Cependant les différentes transitions envisageables ne sont pas toutes réalisables. Les *transitions permises*, celles qui s'accompagnent de l'émission ou de l'absorption d'un photon, sont régies par des *règles* dites *de sélection*, et les *transitions interdites* sont régies par des *règles d'interdiction*. Ces règles ont été établies en spectroscopie par une voie purement empirique et paraissaient quelque peu mystérieuses. Il est vrai aussi que certaines de ces règles ont été interprétées dans la théorie atomique de Bohr sur la base du principe de correspondance. Le développement de la mécanique quantique a permis de tirer au clair la signification physique des règles de sélection. Il apparut que chacune de ces règles exprimait une *loi de la conservation* exacte ou approchée.

2. Les règles de sélection les plus importantes s'appliquant à l'émission ou à l'absorption de la lumière sont des conséquences

de la *loi de la conservation du moment cinétique*. On ne considérera que les *processus à un photon* et on excluera les processus très peu probables de l'émission simultanée de deux ou de plusieurs photons. La loi de la conservation du moment cinétique dans le cas de l'émission d'un photon par l'atome peut s'exprimer sous la forme

$$\boldsymbol{J} = \boldsymbol{J}' + \boldsymbol{s}_{ph}, \tag{39.1}$$

où $\boldsymbol{J}$ représente (en unités $\hbar$) le moment cinétique de l'atome avant émission du photon, $\boldsymbol{J}'$ le moment cinétique après émission et $\boldsymbol{s}_{ph}$ le vecteur spin du photon. Dans ce qui suit on omettra l'indice « ph » pour alléger l'écriture. La loi (39.1) est écrite sous forme symbolique puisque dans un même état les trois composantes du vecteur quantique $\boldsymbol{J}$ ne peuvent avoir des valeurs déterminées. Néanmoins cette forme d'écriture ne peut donner lieu à indétermination dans les développements ultérieurs puisqu'il n'y sera pas question des vecteurs $\boldsymbol{J}$, $\boldsymbol{J}'$ et $\boldsymbol{s}$, mais seulement des nombres quantiques correspondants J, J', s. Il va de soi que les nombres quantiques figurant dans les deux membres de (39.1) doivent être égaux. C'est ce qui sera mis à profit dans ce qui suit en notant que les nombres quantiques figurant dans le second membre de (39.1) s'obtiennent conformément à la règle de composition vectorielle (cf. § 32).

Il existe cependant en mécanique quantique un cas particulier où le vecteur $\boldsymbol{J}$ est déterminé de façon univoque, c'est lorsque le nombre quantique du moment total $J = 0$. On a alors $\boldsymbol{J}^2 = J(J+1) = 0$, ce qui signifie que le vecteur $\boldsymbol{J}$ ainsi que toutes ses projections ont des valeurs déterminées. A ce point de vue le vecteur $\boldsymbol{J}$ se comporte comme en mécanique classique. Il en résulte que les transitions quantiques de l'état quantique caractérisé par $J = 0$ et à un autre état également caractérisé par $J = 0$ (transitions dites 0-0) sont *absolument interdites*. S'il n'en était pas ainsi, le photon ayant un spin, le moment cinétique de l'atome serait différent de zéro au moins dans l'un de ces états, ce qui est contraire à l'hypothèse.

3. La démonstration quantique rigoureuse des règles de sélection exigerait l'introduction de notions et de méthodes mathématiques sortant du cadre de ce cours. On est donc contraint de procéder de façon pas très conséquente en appliquant la *méthode des diagrammes vectoriels*, dont le caractère conventionnel a déjà été mentionné au § 32, 5. Cette méthode ne constitue pas à proprement parler une démonstration, c'est plutôt un procédé mnémonique aidant à saisir la portée des règles de sélection. La seule justification de cette méthode est qu'elle conduit à des résultats corrects. Dans cette méthode les symboles J et s sont assimilés à des *vecteurs classiques* usuels, à cette différence près que leurs modules sont égaux non pas à J et s, mais à $\sqrt{J(J+1)}$ et $\sqrt{s(s+1)}$. (Néanmoins en posant $|J| = J$ et $|s| = s$ on obtient les mêmes règles de sélection.)

La figure 68, *a* exprime la loi de la conservation du moment cinétique lors de l'émission du photon dans le modèle vectoriel considéré: $J = J' + s$.

Considérons d'abord le cas de l'émission d'un photon dans le cas où ni le vecteur J ni le vecteur J' ne sont nuls et $|J'| \geqslant |J|$. N'importe quel côté d'un triangle est plus court que la somme des deux autres côtés. Choisissons le côté le plus long des deux côtés J et J' du triangle, ce qui revient à utiliser l'inégalité $|J'| \leqslant |J| + |s|$ ou

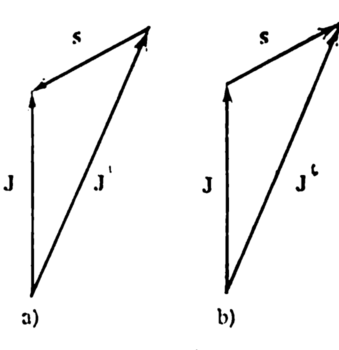


Fig. 68

$$\sqrt{J'(J'+1)} \leqslant \sqrt{J(J+1)} + \sqrt{s(s+1)}. \qquad (39.2)$$

Puisque pour le photon $s = 1$, le dernier terme est égal à $\sqrt{2}$. Les nombres quantiques J et J' sont des nombres entiers lorsque l'atome comporte un nombre pair d'électrons et ils sont demi-entiers si le nombre d'électrons est impair. Il s'ensuit que l'accroissement $\Delta J \equiv J' - J$ peut être soit nul soit égal à un nombre entier positif, puisque lors de l'émission d'un photon le nombre d'électrons contenus dans l'atome ne change pas. En remplaçant dans l'inégalité (39.1) J' par $J + \Delta J$ et en élevant au carré on obtient

$$\Delta J^2 + (2J+1)\Delta J - 2 \leqslant 2\sqrt{2J(J+1)}. \qquad (39.3)$$

Si J a une valeur donnée et $\Delta J \geqslant 0$, le premier membre de cette inégalité croît avec ΔJ, vu que sa dérivée par rapport à ΔJ est essentiellement positive. Pour $\Delta J = 0$ l'inégalité (39.3) est vérifiée. Elle est vérifiée aussi pour $\Delta J = 1$ puisque alors elle se ramène à l'inégalité évidente $J \leqslant \sqrt{2J(J+1)}$. Mais pour $\Delta J = 2$, l'inégalité (39.3) cesse d'être vérifiée puisqu'elle s'écrit alors $2(J+1) \leqslant \sqrt{2J(J+1)}$, résultat erroné comme on peut le constater en élevant au carré cette inégalité. Pour de grandes valeurs de ΔJ l'inégalité (39.3) est à fortiori en défaut.

Le cas $J' \leqslant J$ se ramène au précédent en remplaçant J par J' et vice versa.

Par conséquent, si aucun des nombres quantiques J et J' n'est nul, on arrive à la règle de sélection pour l'émission d'un photon

$$\Delta J \equiv J' - J = \pm 1 \quad \text{ou} \quad 0. \qquad (39.4)$$

Lorsque l'un des nombres J ou J' est égal à zéro, le triangle de la figure 68 dégénère en deux segments de droite égaux ayant même sens ou des sens opposés. Le cas $\Delta J = 0$ se trouvant exclu dans (39.4), les seules transitions possibles correspondent à $\Delta J = \pm 1$.

On a déjà signalé qu'il est impossible que les deux nombres J et J' soient nuls simultanément.

Les règles de sélection pour l'*absorption* d'un photon s'établissent comme pour l'émission, mais comme on a alors $\boldsymbol{J} + \boldsymbol{s} = \boldsymbol{J}'$, on doit remplacer la figure 68, *a* par la figure 68, *b*.

Formulons maintenant les règles de sélection auxquelles doivent satisfaire les nombres quantiques m_J et m'_J des projections du moment cinétique total de l'atome avant et après émission ou absorption d'un photon. Il n'est pas nécessaire d'utiliser ici le modèle vectoriel puisqu'on peut écrire aussitôt

$$\Delta m_J \equiv m'_J - m_J = \pm 1 \quad \text{ou} \quad 0. \tag{39.5}$$

Ces règles doivent être vérifiées en même temps que les règles de sélection précédentes. Dans le cas particulier où les projections m_J et m'_J prennent leurs valeurs maximales, elles se confondent avec J et J' et les règles (39.5) se ramènent aux règles (39.4). Mais sont également possibles les situations où l'une, au moins, de ces projections est plus petite que le nombre quantique J correspondant.

4. A cette étape de l'exposé, une remarque s'impose. Il a été indiqué au § 37 que le spin du photon ne peut prendre que deux orientations suivant la direction de sa propagation. Cela implique que n'importe quel état de polarisation du photon peut être réalisé par combinaison linéaire de deux états, l'un à polarisation droite, l'autre à polarisation gauche. D'autre part, s étant donné, le nombre d'états se distinguant par les projections du vecteur $\boldsymbol{s}$ sur une direction donnée devrait être égal à $2s + 1$. On pourrait en déduire que le spin du photon devrait être égal à $^1/_2$. Or, dans ce cas, l'émission ou l'absorption d'un photon devrait donner lieu à une variation du nombre quantique J du moment cinétique total de l'enveloppe électronique de l'atome égale à $\pm ^1/_2$. Cela imposerait que le nombre J d'entier deviendrait demi-entier et inversement. Or cela serait en contradiction avec le fait établi qu'à la suite de l'émission ou de l'absorption d'un photon, le nombre d'électrons ne change pas et que *le nombre quantique J est toujours entier pour un nombre pair d'électrons et demi-entier si le nombre d'électrons est impair*. On avait indiqué au § 37, 8 que pour le photon avec $s = 1$, parmi les trois projections possibles du spin, l'une d'elles est irréalisable par suite de la transversalité des ondes électromagnétiques.

5. Les règles de sélection qui viennent d'être établies pour les processus à un photon sont fondées sur l'application rigoureuse de la loi de la conservation du moment cinétique. Il convient d'établir les règles de sélection qui sont liées au comportement des vecteurs $\boldsymbol{L}$ et $\boldsymbol{S}$. L'émission des ondes électromagnétiques est déterminée par les *propriétés électromagnétiques de l'électron*, i.e. par sa charge et par son moment magnétique. L'émission d'un photon résulte soit d'un changement du mouvement de la charge (variation du vecteur

L), soit d'une rotation du moment magnétique de spin S, soit encore de ces deux processus à la fois. Il est bien évident que l'émission d'un rayonnement déterminée par la rotation du spin est un effet essentiellement relativiste. La théorie montre que dans le cas de l'émission d'une lumière visible l'interaction du photon avec la charge de l'électron est un effet de plusieurs ordres de grandeur plus intense que celui déterminé par l'interaction du photon avec le moment magnétique de l'électron. On en conclut que l'émission d'un photon dans la gamme optique n'est pas liée à une variation de S, i.e.

$$\Delta S = 0. \tag{39.6}$$

Il s'ensuit que l'émission et l'absorption d'une lumière dont la longueur d'onde n'est pas très courte se produisent comme si le spin n'existait pas et que le moment magnétique de l'atome était purement orbital. C'est pour cela qu'on peut utiliser tous les résultats qui ont été obtenus plus haut en y remplaçant le moment total J par le moment orbital L. Par conséquent, pour les processus d'émission et d'absorption d'un seul photon de longueur d'onde pas trop courte, on peut appliquer les règles de sélection approchées

$$\Delta L \equiv L' - L = \pm 1 \quad \text{ou} \quad 0. \tag{39.7}$$

Si L ou L' s'annule, la valeur $\Delta L = 0$ doit être exclue. La valeur $\Delta L = 0$ est exclue aussi pour les atomes ne possédant qu'un *seul électron de valence*, par exemple pour les atomes d'hydrogène ou des métaux alcalins. Cette interdiction est cependant liée non pas à la loi de la conservation du moment cinétique, mais à la *loi de la conservation de la parité de la fonction d'onde*. Cette question sera examinée en Physique nucléaire, et ici on notera seulement que la règle de sélection $\Delta L = \pm 1$ a déjà été utilisée au § 34 pour expliquer les séries spectrales des métaux alcalins.

6. Lorsque $\Delta J = \pm 1$, le photon émis est polarisé *circulairement* et lorsque $\Delta J = 0$, le photon est polarisé *rectilignement*. Ces résultats peuvent paraître surprenants sachant que le spin du photon est égal à l'unité. La mécanique quantique trouve une solution originale à cet obstacle, en affirmant que le photon émis dans ces conditions possède un *spin indéterminé*, quoique cet état résulte de la superposition de deux états à polarisations circulaires gauche et droite, ayant des probabilités égales. Si on mesure le moment cinétique que transmet le photon au corps qui l'absorbe, on peut obtenir avec la même probabilité les résultats $+1$ et -1.

Remarquons avec force que toutes les règles de sélection qui ont été démontrées sont liées aux propriétés du photon et se rapportent aux transitions quantiques donnant lieu à l'émission ou à l'absorption d'un seul photon. Ces règles ne peuvent être étendues ni aux processus d'émission ou d'absorption de plusieurs photons, ni aux

transitions quantiques qui sont déterminées non par le rayonnement électromagnétique mais par d'autres causes, tels les chocs électroniques dans les décharges gazeuses, l'excitation thermique des atomes, etc.

Les transitions radiatives mettant en défaut les règles de sélection sont cependant possibles et sont appelées *transitions interdites*. Leur probabilité est beaucoup plus petite que celle des transitions permises et l'intensité des *raies spectrales interdites* est en général beaucoup plus petite que celle des raies permises.

§ 40. Structure fine des raies spectrales de l'hydrogène et des métaux alcalins

1. Il a été indiqué au § 38 que les termes spectraux éclatent par suite de l'interaction spin-orbite et de la variation de la masse de l'électron avec sa vitesse. Cet éclatement est appelé *structure fine* des termes spectraux ou des niveaux énergétiques. Dans ce paragraphe on ne considérera que les atomes ayant un *seul* électron de valence. En conséquence les termes spectraux seront dénotés par des lettres minuscules. L'énergie des niveaux dépend des nombres quantiques n, l, j, mais en l'absence de champs électrique ou magnétique elle ne peut dépendre du nombre quantique magnétique m_j, puisque alors toutes les directions spatiales sont équivalentes. Ce n'est que dans le cas de l'hydrogène et d'atomes hydrogénoïdes que se réalise une *dégénérescence aléatoire* par rapport à l; celle-ci résulte de ce que le champ électrique du noyau, dans lequel se meut l'unique électron de l'atome, est *coulombien*. Dans ce cas l'énergie du niveau ne dépend que des nombres quantiques n et j et ne dépend plus de l. Sa valeur est donnée par la formule (38.4).

On doit faire la différence entre la *structure fine des termes* et la *structure fine des raies spectrales*, i.e. l'éclatement d'une raie spectrale en plusieurs composantes rapprochées. Cet éclatement est déterminé par les transitions permises entre les différents sous-niveaux formés à partir de certains niveaux énergétiques, donc par les règles de sélection (39.4) et (39.7). Pour les atomes monoélectroniques et en l'absence de champs extérieurs, ces règles se réduisent aux expressions

$$\Delta l = \pm 1 \tag{40.1}$$

et

$$\Delta j = \pm 1 \quad \text{ou} \quad 0. \tag{40.2}$$

2. Examinons à titre d'exemple la structure fine de la raie L_α de la série de Lyman de l'hydrogène:

$$\bar{\nu} = 1s - 2p,$$

où $\bar{\nu} = 1/\lambda$ dénote le nombre d'onde spectroscopique (cf. § 11). Le terme $1s$ est un singulet, le terme $2p$ est un doublet comportant les sous-termes $2p_{1/2}$ et $2p_{3/2}$ (fig. 69). Comme les règles de sélection (40.1) et (40.2) autorisent les transitions des niveaux $2p_{1/2}$ et $2p_{3/2}$ sur le niveau $1s_{1/2}$, la raie L_α doit être un doublet formé de deux raies spectrales rapprochées. La raie $1s_{1/2} \to 2p_{3/2}$ est plus intense que la raie $1s_{1/2} \to 2p_{1/2}$; la raie $1s_{1/2} \to 2p_{3/2}$ et la transition correspondante $2p_{3/2} \to 1s_{1/2}$ sont représentées sur la figure 69

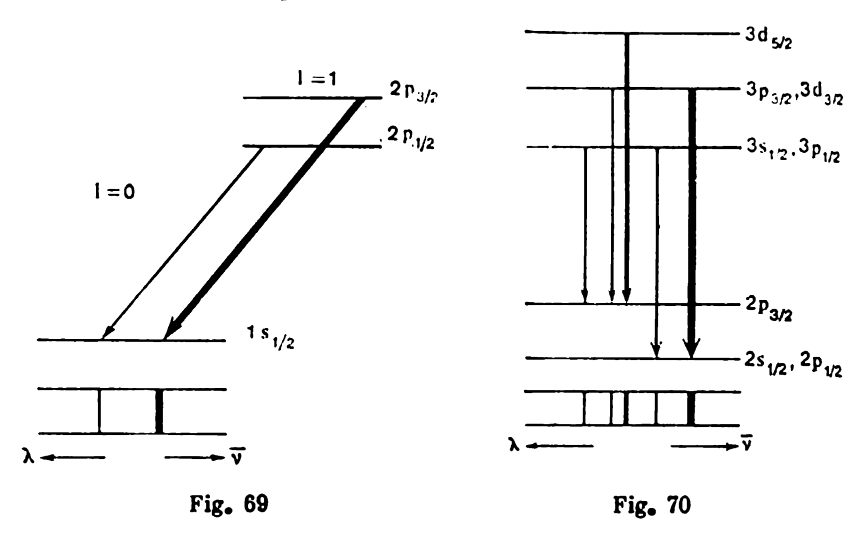


Fig. 69 Fig. 70

par des traits plus épais. La distance de séparation des deux raies évaluée à l'aide de la formule (38.4), est égale à $\Delta\bar{\nu} = 0{,}365$ cm^{-1} et en longueurs d'onde elle vaut $\Delta\lambda = 5{,}3 \cdot 10^{-4}$ nm. La longueur d'onde de la raie L_α est $\lambda = 121{,}6$ nm. Pour pouvoir séparer la structure fine de la raie L_α l'appareil spectral doit avoir un pouvoir de résolution égal ou supérieur à $\lambda/\Delta\lambda = 121{,}6/(5{,}3 \cdot 10^{-4}) \approx$ $\approx 2{,}23 \cdot 10^5$. Les autres raies de la série de Lyman de l'hydrogène: L_β, L_γ, L_δ ont aussi une structure de doublets analogue.

Un deuxième exemple concerne la structure fine de la raie H_α de la série de Balmer de l'hydrogène (fig. 70). Au nombre quantique principal $n = 2$ correspondent les états $2s_{1/2}$, $2p_{1/2}$ et $2p_{3/2}$. Comme l'énergie ne dépend pas du nombre orbital l, les deux premiers états ont la même énergie, et par suite sont représentés par un niveau énergétique formé de deux sous-niveaux confondus. Ainsi, avec $n = 2$ on obtient deux niveaux, dont un est formé de deux niveaux confondus. Avec $n = 3$ on obtient trois niveaux: un niveau simple $3d_{5/2}$ et deux niveaux ($3p_{3/2}$, $3d_{3/2}$) et ($3s_{1/2}$, $3p_{1/2}$) qui sont confondus. La raie H_α apparaît à la suite de transitions quantiques des

niveaux supérieurs ($n = 3$) sur les niveaux inférieurs ($n = 2$). On a représenté sur la figure 70 les transitions satisfaisant aux règles de sélection (40.1) et (40.2) ainsi que les raies spectrales correspondantes avec indication de leurs intensités relatives approchées. La raie H_α ainsi que les autres raies de la série de Balmer H_β, H_γ, H_δ, . . . comportent chacune cinq composantes.

3. La structure fine des niveaux et des raies spectrales des métaux alcalins et des ions qui leur ressemblent est déterminée surtout par les *interactions spin-orbite* et non pas par la dépendance de la

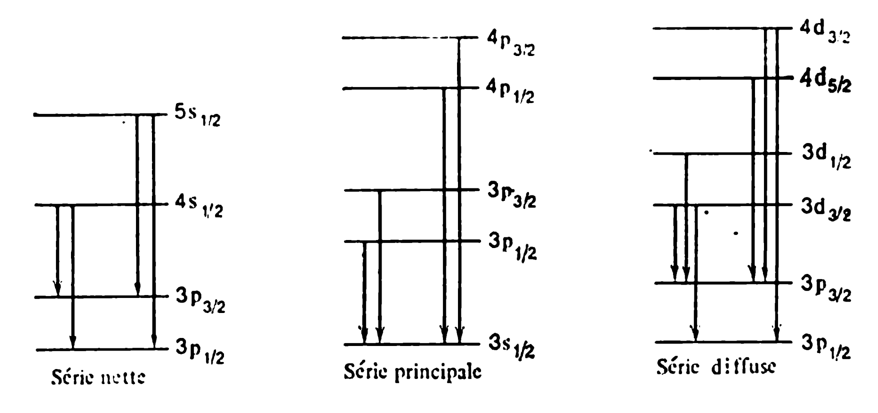


Fig. 71

masse de l'électron avec la vitesse. La raison en est que les éclatements dus aux interactions spin-orbite s'intensifient à mesure qu'augmente Z. Lorsque Z est grand l'éclatement des niveaux dû à la variation de la masse de l'électron avec la vitesse est négligeable devant l'éclatement déterminé par les interactions spin-orbite. Le tableau des éclatements fins est plus simple pour les métaux alcalins que pour l'hydrogène. Pour ce dernier la situation est rendue compliquée par la dégénérescence par rapport à l. Cette dégénérescence est levée chez les métaux alcalins et les régularités de l'éclatement fin se présentent de façon parfaitement nette.

La figure 71 illustre les causes d'apparition de la structure fine dans les séries spectrales du sodium. La série principale a pour origine les transitions des différents niveaux p sur le niveau profond $3s_{1/2}$. Ce dernier est un niveau simple tandis que tous les niveaux p sont des doublets, dont les intervalles de séparation décroissent à mesure qu'augmente le nombre quantique principal n. C'est pour cela que toutes les raies spectrales de la série principale sont des *doublets*. Les distances de séparation des composantes des doublets diminuent à mesure qu'augmentent leurs numéros d'ordre (i.e. à mesure qu'augmente la fréquence). La série principale apparaît

aussi dans les spectres d'absorption puisqu'à l'état normal les atomes de sodium occupent le niveau le plus bas, i.e. le niveau $3s_{1/2}$. La raie la plus intense est la raie de résonance jaune du sodium qui résulte des transitions $3p_{1/2} \to 3s_{1/2}$ et $3p_{3/2} \to 3s_{1/2}$ auxquelles correspondent les longueurs d'onde $\lambda_1 = 589,6$ nm et $\lambda_2 = 589,0$ nm, la différence $\Delta\lambda = 0,6$ nm. Notons que Z augmentant l'éclatement des raies spectrales des métaux alcalins devient toujours plus important. Aussi la raie de résonance du rubidium est formée de deux composantes dont les longueurs d'onde sont: $\lambda_1 = 794.8$ nm et $\lambda_2 = 780,0$ nm, l'intervalle étant égal à $\Delta\lambda = 14,8$ nm. Pour le césium on a $\lambda_1 = 894,4$ nm et $\lambda_2 = 852,1$ nm, i.e. $\Delta\lambda = 42,3$ nm. Pour des dédoublements aussi importants le terme « structure fine » semble peu adéquat.

Les raies de la série nette sont aussi des *doublets* et résultent du passage de niveaux s simples au niveau double $3p$ situé plus bas et composé des sous-niveaux $3p_{1/2}$ et $3p_{3/2}$. Aussi les intervalles entre les composantes des doublets sont les mêmes pour toute la série, et toutes les composantes sont des raies *nettes*. C'est pour cette raison que toute la série a été appelée série *nette*.

La série diffuse résulte de transitions des niveaux supérieurs doubles ($3d_{3/2}$, $3d_{1/2}$) et ($4d_{5/2}$, $4d_{3/2}$), etc., sur le niveau inférieur double ($3p_{1/2}$, $3p_{3/2}$). Ses raies spectrales sont des *triplets*, puisque les transitions telles que $3d_{5/2} \to 3p_{1/2}$, où le nombre j varie de deux unités, sont interdites conformément aux règles de sélection (40.2). Les distances entre les sous-niveaux des différents niveaux d sont beaucoup plus petites que celles des sous-niveaux des niveaux p. En conséquence, si l'appareil spectral a un pouvoir de résolution faible, les composantes des triplets ne peuvent être séparées et la raie apparaît floue. C'est pour cela que la série est appelée série *diffuse*.

PROBLÈMES

Calculer la structure fine de la raie spectrale H_α de la série de Balmer de l'hydrogène.

Solution. Les niveaux énergétiques de l'atome d'hydrogène sont définis par la formule (38.4), où on pose $Z = 1$, i.e.

$$\mathscr{E}_{n,j} = -\frac{\mu e^4}{2\hbar^2 n^2}\left[1 + \frac{\alpha^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right)\right]. \qquad (40.3)$$

En y remplaçant la masse μ de l'électron par la masse réduite, récrivons cette formule sous la forme

$$\mathscr{E}_{n,j} = -\frac{R_H}{n^2}\left[1 + \frac{\alpha^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right)\right], \qquad (40.3a)$$

où R_H est la constante de Rydberg pour l'hydrogène définie par les expressions (13.10) et (13.9):

$$R_H = 109677,576 \text{ cm}^{-1}.$$

Les niveaux énergétiques sont mesurés ici en cm^{-1}. Le carré de la constante de structure fine est égal à $\alpha^2 = 5{,}3251 \cdot 10^{-5}$. L'origine de la raie H_α est illustrée par la figure 70. On en tire

$$\mathscr{E}_{2,1/2} = -27419{,}395 \cdot 1{,}00001664 = -27419{,}850\ cm^{-1},$$
$$\mathscr{E}_{2,3/2} = -27419{,}395 \cdot 1{,}00000333 = -27419{,}485\ cm^{-1},$$
$$\mathscr{E}_{3,1/2} = -12186{,}397 \cdot 1{,}0000197 = -12186{,}639\ cm^{-1},$$
$$\mathscr{E}_{3,3/2} = -12186{,}397 \cdot 1{,}0000067 = -12186{,}478\ cm^{-1},$$
$$\mathscr{E}_{3,5/2} = -12186{,}397 \cdot 1{,}0000022 = -12186{,}423\ cm^{-1}.$$

En calculant les différences de ces nombres conformément à la figure 70, on trouve les nombres d'onde spectroscopiques (en cm^{-1}) et les longueurs d'onde (en nm) des cinq composantes produites par l'éclatement de la raie H_α de l'hydrogène :

	$\bar{\nu}$, cm^{-1}	λ, nm
1.	15233,372	656,4534
2.	15233,211	656,4604
3.	15233,067	656,4660
4.	15233,007	656,4692
5.	15232,846	656,4761

On notera que les longueurs d'onde et les nombres d'onde spectroscopiques sont rapportés au vide (et non pas à l'air, comme il est d'usage en spectroscopie).

§ 41. Effets Zeeman normal et anormal

1. Dans un champ coulombien (dans l'approximation non relativiste) tous les niveaux énergétiques de l'électron sont *dégénérés*, i.e. l'énergie ne dépend que du nombre quantique principal n, mais ne dépend pas du nombre orbital l. C'est pour cela que toutes les raies spectrales de l'atome d'hydrogène sont *simples* (singulets). On peut considérer les atomes des métaux alcalins comme des atomes monoélectroniques, où l'électron se meut dans un champ à symétrie centrale, qui *n'est* cependant *pas un champ coulombien*. La dégénérescence sur l est alors levée et l'énergie des niveaux dépend maintenant de n et de l. C'est à cette circonstance qu'est liée l'apparition des séries spectrales des métaux alcalins. L'existence d'une interaction spin-orbite détermine la *structure fine* des raies spectrales, mais en l'absence de champs extérieurs toutes les directions spatiales sont équivalentes et par suite les énergies des niveaux *ne dépendent pas du nombre quantique magnétique* m_J, quoique pour un J donné, le nombre m_J peut prendre $2J + 1$ valeurs. La multiplicité de la dégénérescence correspondante est donc égale à $2J + 1$. *L'application d'un champ magnétique lève cette dégénérescence* et chaque niveau

énergétique éclate en donnant $2J + 1$ sous-niveaux. C'est ce qui explique l'*effet Zeeman*, qui a fait l'objet d'une étude classique détaillée dans le t. IV, § 92. Or avant que fut découvert le spin, l'application des règles de sélection ne permettait à la théorie quantique d'expliquer que l'*effet Zeeman simple.*

En effet, si l'atome possède un moment magnétique $\vec{\mathfrak{m}}$, son énergie dans un champ magnétique B sera égale à $\mathscr{E} = \mathscr{E}_0 - (\vec{\mathfrak{m}}B)$, où $\mathscr{E}_0$ est l'énergie dans un champ magnétique nul. Si le spin n'existe pas, le moment magnétique ne peut être déterminé que par le mouvement orbital des électrons, et sa projection sur la direction du champ magnétique sera égale à un nombre entier de magnétons de Bohr, i.e. $m_L \mathfrak{m}_B = m_L (e\hbar/2\mu c)$. Par conséquent,

$$\mathscr{E} = \mathscr{E}_0 - m_L (e\hbar B/2\mu c) = \mathscr{E}_0 - \hbar\Omega m_L, \tag{41.1}$$

où

$$\Omega = \frac{eB}{2\mu_e c} \tag{41.2}$$

est la *fréquence de Larmor*. Dans un champ magnétique chaque niveau se scinde en $2L + 1$ sous-niveaux. Les transitions quantiques entre les différents niveaux donnent naissance à des raies spectrales de fréquences

$$\omega = \omega_0 - \Omega \Delta m_L,$$

où $\omega_0 \equiv \Delta\mathscr{E}_0/\hbar$ est la fréquence de la raie émise en champ magnétique nul. En vertu des règles de sélection $\Delta m_L = 0$ ou ± 1. Par conséquent, la fréquence de la raie émise sera égale à

$$\omega = \omega_0 \quad \text{ou bien} \quad \omega_0 \pm \Omega, \tag{41.3}$$

ce qui correspond au *triplet de Lorentz*. Selon la théorie classique aux transitions avec $\Delta m_L = 0$ correspondent des vibrations parallèles à la direction du champ magnétique (composantes π) et aux transitions avec $\Delta m_L = \pm 1$ des vibrations transversales par rapport au champ magnétique (composantes σ).

2. La prise en considération du *spin de l'électron* permet d'expliquer l'*effet Zeeman complexe*. Supposons qu'on a affaire à une *liaison normale* (liaison de Russel-Saunders, cf. § 38, 3) et utilisons le modèle vectoriel. Dans ce cas l'atome est caractérisé par le moment cinétique orbital L, par le moment de spin S et par le moment cinétique total $J = L + S$. Sous forme opératorielle $\hat{J} = \hat{L} + \hat{S}$, $\hat{J}^2 = \hat{L}^2 + \hat{S}^2 + 2(\hat{L}\hat{S})$. Considérons un état pour lequel les carrés des moments ont des valeurs déterminées et sont respectivement égaux à $J(J + 1)$, $L(L + 1)$ et $S(S + 1)$. Pour ce même état le produit scalaire

(LS) a aussi une valeur déterminée égale à

$$(LS) = \frac{J^2 - L^2 - S^2}{2} = \frac{J(J+1) - L(L+1) - S(S+1)}{2}. \quad (41.4)$$

Aux moments cinétiques L et S correspondent le moment magnétique orbital $\vec{\mathfrak{m}}_L = -g_l L$ et le moment magnétique de spin $\vec{\mathfrak{m}}_s = -g_s S$. (Le signe moins est utilisé pour la raison que la charge de l'électron est négative et que par suite les vecteurs $\vec{\mathfrak{m}}_L$ et L et les vecteurs $\vec{\mathfrak{m}}_s$ et S sont orientés en sens inverses.) Prenons $\hbar$ pour unité de moment cinétique et le magnéton de Bohr $\mathfrak{m}_B$ pour unité de moment magnétique. Avec ces unités on a pour l'électron $g_l = 1$, $g_s = 2$. Le fait que $g_l \neq g_s$ détermine l'existence de l'effet Zeeman anormal. Pour l'instant on s'abstiendra de préciser les valeurs numériques de g_l et g_s vu que les valeurs de g_s pour le proton et le neutron ne sont pas égales à la valeur de g_s pour l'électron.

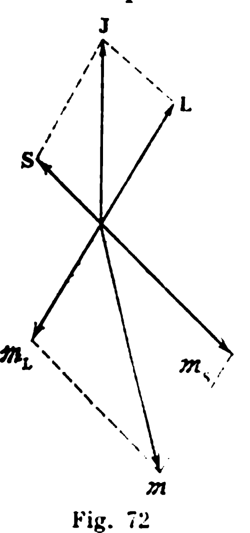


Fig. 72

En l'absence de champ extérieur le moment cinétique total J se conserve en module et en direction (si on utilise le modèle vectoriel). Les moments L et S ne se conservent pas par suite de l'interaction spin-orbite, mais comme dans l'état considéré leurs modules se conservent ces moments exécutent une *précession régulière* autour de la direction fixe du vecteur J, et ce avec la même vitesse angulaire puisque les vecteurs L, S, J doivent être toujours contenus dans un même plan. Les moments magnétiques $\vec{\mathfrak{m}}_L = -g_l L$ et $\vec{\mathfrak{m}}_s = -g_s S$, ainsi que le moment magnétique total $\vec{\mathfrak{m}} = -g_l L - g_s S$ exécutent une précession avec la même vitesse angulaire. En effet, puisque g_l et g_s sont inégaux le vecteur $\vec{\mathfrak{m}}$ n'est pas colinéaire au vecteur J et sa direction doit donc varier (voir la figure 72 où il a été tenu compte de ce que vu la charge négative de l'électron les vecteurs L et $\vec{\mathfrak{m}}_L$, et les vecteurs S et $\vec{\mathfrak{m}}_s$ sont antiparallèles).

Déterminons maintenant la projection $\vec{\mathfrak{m}}_{\parallel}$ du vecteur $\vec{\mathfrak{m}}$ sur la direction du vecteur J. Pour cela on doit calculer d'abord le produit scalaire

$$(\vec{\mathfrak{m}}J) = (-g_l L - g_s S)(L + S) = -g_l L^2 - g_s S^2 - (g_l + g_s)(LS),$$

ce qui en vertu de (41.4) se laisse écrire

$$(\vec{\mathfrak{m}}J) = -gJ^2, \quad (41.5)$$

où g dénote la quantité

$$g=\frac{g_l+g_s}{2}+\frac{g_l-g_s}{2}\frac{L^2-S^2}{J^2}$$

ou encore

$$g=\frac{g_l+g_s}{2}+\frac{g_s-g_l}{2}\frac{S(S+1)-L(L+1)}{J(J+1)}. \qquad (41.6)$$

En particulier, pour l'électron $g_l = 1$, $g_s = 2$ et l'expression (41.6) s'écrit

$$g=\frac{3}{2}+\frac{S(S+1)-L(L+1)}{2J(J+1)}. \qquad (41.7)$$

Dans ce cas g et une *fraction rationnelle*. La quantité g est appelée *facteur de Landé* (1888-1975).

Ainsi, en vertu de (41.5) on peut écrire $\mathfrak{m}_\parallel |J| = -g|J|^2$, ce qui montre que la projection du vecteur $\mathfrak{m}$ sur la direction du vecteur J a une valeur déterminée $\mathfrak{m}_\parallel = -g|J|$. La projection orthogonale $\mathfrak{m}_\perp$ n'a pas de valeur déterminée. Dans le modèle vectoriel cette projection exécute une précession autour du vecteur J. Lorsqu'on envisage des processus qui sont lents devant cette précession, on peut négliger cette composante orthogonale et poser que le moment magnétique total de l'atome se réduit à sa projection $\mathfrak{m}_\parallel$. Dans cette approximation

$$\mathfrak{m} = -\mathfrak{m}_\parallel = -gJ. \qquad (41.8)$$

C'est ce moment magnétique qui se manifeste dans les expériences de Stern et Gerlach.

Les formules (41.6) et (41.7) peuvent être obtenues par un raisonnement quantique conséquent. Il suffit d'introduire l'*opérateur projection du moment magnétique* $\hat{\mathfrak{m}}_\parallel$ et de déterminer ses valeurs propres. On n'exposera pas ce procédé, parce que la démonstration qui vient d'être donnée, bien que n'étant pas conséquente, a pour elle l'avantage d'être concrète. Notons qu'il est aisé d'appliquer cette méthode concrète au cas de la liaison $(j\text{-}j)$ ainsi qu'aux liaisons intermédiaires. On obtient alors pour g des formules qui diffèrent de (41.6) et (41.7) et de ce fait le caractère de l'éclatement des raies spectrales dans l'effet Zeeman anormal est différent de celui qui correspond au cas de la liaison normale. Le lecteur que cette question intéresserait pourrait trouver des développements complets dans l'ouvrage de M. A. Eliachevitch (né en 1908), *Spectroscopie atomique et moléculaire* (en russe), Moscou, 1962.

3. On a supposé ci-dessus que le champ magnétique était nul. Considérons maintenant le cas où l'atome est soumis à l'action d'un champ magnétique homogène B. Le moment total J de l'atome ne se conserve alors plus, et se met à exécuter sous l'action de ce champ une *précession forcée*. Examinons d'abord le cas d'un champ faible

lorsque la précession qu'il détermine est lente par rapport à la précession déterminée par l'interaction spin-orbite. Cela signifie que le champ B doit être suffissamment faible pour que l'éclatement des raies spectrales par effet Zeeman qu'il provoque soit petit devant l'écartement des composantes de la structure fine des raies spectrales. Dans ces conditions on observe l'effet Zeeman *anormal*. Cette condition ne peut jamais être satisfaite pour les raies simples (singulets), puisque pour ces raies n'importe quel champ B est un champ *fort* et l'effet Zeeman qui se manifeste est l'effet normal.

Dans le cas d'un champ magnétique faible, l'énergie supplémentaire qu'un atome y acquiert peut être calculée en supposant que son moment magnétique total est donné par l'expression (41.8), ou en unités usuelles par l'expression

$$\vec{\mathfrak{m}} = -g\frac{e\hbar}{2\mu c}J. \qquad (41.9)$$

Par conséquent,

$$\mathscr{E} = \mathscr{E}_0 - (\vec{\mathfrak{m}}B) = \mathscr{E}_0 + gm_J\Omega\hbar, \qquad (41.10)$$

où $\mathscr{E}_0$ est l'énergie de l'atome dans un champ magnétique nul. Cette formule indique en combien de sous-niveaux se scinde chaque niveau énergétique de l'atome lorsqu'on le soumet à l'action d'un champ magnétique faible. Cette même formule détermine, compte tenu des règles de sélection, les transitions radiatives possibles entre les niveaux et les longueurs d'ondes des raies spectrales correspondantes. De toute évidence on ne doit conserver que la règle de sélection suivante:

$$\Delta m_J = \pm 1 \quad \text{ou bien} \quad 0 \qquad (41.11)$$

puisqu'il est nécessaire que la raie spectrale dont on envisage l'éclatement Zeeman ne soit pas interdite et satisfasse donc à toutes les autres règles de sélection. On peut donc dire que les formules (41.10) et (41.11) servent de fondement à l'explication de l'effet Zeeman normal et anormal.

4. Pour mettre en évidence l'origine de l'effet Zeeman anormal examinons la raie spectrale d'un métal alcalin correspondant à la transition du niveau doublet ($^2p_{1/2}$, $^2p_{3/2}$) au niveau singulet $^2s_{1/2}$. Dans le cas où $n = 3$ on obtient la raie double jaune du sodium. On s'abstiendra de fixer la valeur du nombre quantique principal pour bien montrer que l'effet Zeeman anormal n'en dépend pas. Evaluons d'abord le facteur de Landé par la formule (41.7). Pour le niveau $p_{3/2}$, $l = 1$, $j = 3/2$, $s = 1/2$ on trouve $g = 4/3$. De même pour le niveau $p_{1/2}$, $g = 2/3$ et pour le niveau $s_{1/2}$, $g = 2$. Selon la formule (41.10) les déplacements des sous-termes, dérivant des termes considérés, sont égaux à gm_J, à condition de prendre pour unité l'éclatement d'après la théorie de Lorentz, i.e. la fréquence de

Larmor Ω. Les valeurs ainsi obtenues sont consignées dans le tableau suivant :

Tableau 3

$p_{3/2}$	m_j	$-3/2$	$-1/2$	$+1/2$	$+3/2$
	gm_j	-2	$-2/3$	$+2/3$	$+2$
$p_{1/2}$	m_j		$-1/2$	$+1/2$	
	gm_j		$-1/3$	$+1/3$	
$s_{1/2}$	m_j		$-1/2$	$+1/2$	
	gm_j		-1	$+1$	

Les résultats du calcul de l'éclatement du doublet $p_{1/2} \rightarrow s_{1/2}$; $p_{3/2} \rightarrow s_{1/2}$ dans un champ magnétique faible sont consignés dans le tableau 4, où les raies sont numérotées dans l'ordre des fréquences

Tableau 4

Transitions $m_j^{(1)} \rightarrow m_j^{(2)}$	Polarisation	$(gm_j)^1 \rightarrow (gm_j)^2$	n° de la raie
	Eclatement de la raie $p_{1/2} \rightarrow s_{1/2}$		
$-1/2 \rightarrow +1/2$	σ	$-1/3-1=-4/3$	1
$-1/2 \rightarrow -1/2$	π	$-1/3-(-1)=2/3$	3
$+1/2 \rightarrow +1/2$	π	$1/3-1=-2/3$	2
$+1/2 \rightarrow -1/2$	σ	$1/3-(-1)=4/3$	4
	Eclatement de la raie $p_{3/2} \rightarrow s_{1/2}$		
$-3/2 \rightarrow +1/2$	interdite		
$-3/2 \rightarrow -1/2$	σ	$-2-(-1)=-1$	6
$-1/2 \rightarrow +1/2$	σ	$-2/3-1=-5/3$	5
$-1/2 \rightarrow -1/2$	π	$-2/3+1=1/3$	8
$+1/2 \rightarrow +1/2$	π	$2/3-1=-1/3$	7
$+1/2 \rightarrow -1/2$	σ	$2/3+1=5/3$	10
$+3/2 \rightarrow +1/2$	σ	$2-1=1$	9
$+3/2 \rightarrow -1/2$	interdite		

croissantes. A ce tableau est adjoint un schéma des transitions spectrales (fig. 73). Sur cette figure on a indiqué dans la colonne de gauche les niveaux énergétiques $s_{1/2}$, $p_{1/2}$, $p_{3/2}$ pour un champ magnétique B nul, et à droite de cette colonne on a indiqué l'éclatement de ces mêmes niveaux dans un champ magnétique B faible, ainsi que toutes les transitions permises entre ces niveaux caractérisant l'effet Zeeman dans le cas considéré. On voit que la raie $p_{1/2} \to s_{1/2}$ se scinde en quatre composantes, les extrêmes étant des composantes σ et les centrales des composantes π. La deuxième raie $p_{3/2} \to s_{1/2}$ se scinde en six composantes: deux composantes π au milieu, entourées de quatre composantes σ. Les distances entre les composantes sont données en unités d'éclatement de Lorentz, i.e. pour unité de fréquence on a pris la fréquence de Larmor $eB/2\mu c$.

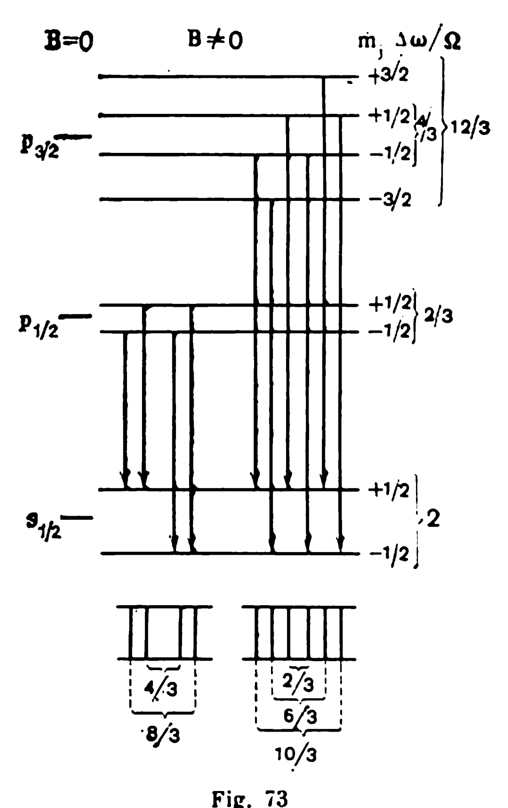


Fig. 73

5. Les quantités gm_j figurant dans (41.10) sont de petits nombres rationnels. C'est pour cela qu'à l'aide du schéma expliquant l'apparition de l'effet Zeeman anormal on arrive à la règle qui fut établie empiriquement par Runge (1856-1927) à une époque où la nature de l'effet Zeeman anormal était encore inconnue. La *règle de Runge* affirme que les intervalles entre les composantes apparaissant dans l'effet Zeeman anormal par éclatement dans un champ magnétique s'expriment par de *petits nombres rationnels*, si on adopte pour unité l'intervalle d'éclatement normal de Lorentz Ω, observé dans le même champ magnétique.

Il existe une deuxième règle empirique qui fut également établie avant que fut élaborée la théorie de l'effet Zeeman anormal. Il s'agit de la *règle de Preston*, selon laquelle les raies spectrales caractérisées par le *même symbole de série*, présentent le *même type d'éclatement dans un champ magnétique*, quelles que soient les valeurs des nom-

bres quantiques principaux. On entend par symbole de série les expressions de la forme ${}^2P_{3/2}{}^2D_{5/2}$, où ${}^2D_{5/2}$ caractérise l'état de départ et ${}^2P_{3/2}$ l'état d'arrivée des transitions radiatives. Les valeurs des nombres quantiques principaux dans ces états de l'atome peuvent être quelconques. Par exemple, selon la règle de Preston les composantes ${}^2S_{1/2}{}^2P_{1/2}$ des doublets des séries principales des métaux alcalins éclatent également dans un champ magnétique faible indépendamment de la valeur du nombre quantique principal et de la nature de l'élément chimique considéré. La même conclusion se rapporte aux secondes composantes des mêmes doublets ainsi qu'aux multiplets des séries nette et diffuse des métaux alcalins. Les écarts par rapport à la règle de Preston sont liés soit à l'étroitesse du multiplet, lorsque les distances entre ses composantes sont inférieures ou comparables à l'éclatement déterminé par l'application du champ magnétique, soit à une liaison anormale entre L et S.

6. Lorsque l'intensité du champ magnétique augmente et que l'action qu'exerce celui-ci sur l'atome devient égale ou supérieure à l'interaction spin-orbite, le caractère de l'éclatement des raies par effet Zeeman devient autre. Dans ces conditions on ne peut plus étudier séparément le comportement de chaque composante du multiplet de la structure fine et on doit considérer le comportement du *multiplet en entier.* A mesure que le champ devient plus fort le multiplet se modifie. On peut s'attendre à des résultats particulièrement simples lorsque le champ magnétique devient assez fort pour que la largeur de l'éclatement devienne notablement plus grande que la largeur du multiplet initial. Dans ces conditions on peut négliger l'interaction spin-orbite et étudier de façon *indépendante* la précession forcée des vecteurs $\boldsymbol{L}$ et $\boldsymbol{S}$ dans le champ magnétique. On a alors

$$\vec{\mathfrak{m}} = \vec{\mathfrak{m}}_L + \vec{\mathfrak{m}}_S = -\frac{e\hbar}{2\mu c}(L+2S)$$

et par conséquent

$$\mathscr{E} = \mathscr{E}_0 + \frac{e\hbar}{2\mu c}(L+2S)\,\boldsymbol{B} = \mathscr{E}_0 + \Omega\hbar\,(m_L + 2m_S).$$

D'après les règles de sélection $\Delta m_L = \pm 1$ ou 0 et $\Delta m_S = 0$; par conséquent

$$\Delta\omega = \Delta\mathscr{E}/\hbar = \pm\,\Omega \quad \text{ou} \quad 0.$$

Effectivement, Paschen (1865-1947) et Back (1881-1959) découvrirent en 1912 qu'à mesure que le champ magnétique augmentait commençaient à se manifester des transformations magnéto-optiques qui *aboutissent dans les champs forts à un triplet simple de Zeeman-Lorentz*, quoique chacune de ses trois composantes comporte encore des éclatements indépendants de l'intensité du champ magnétique qui sont du même ordre que la structure fine des multiplets en

champ nul et qui sont petits devant les intervalles entre les composantes de triplet. Cet effet est connu sous le nom d'*effet Paschen-Back*.

7. La figure 74 illustre les résultats que l'on obtient si on photographie en cours d'expérience suivant une direction perpendiculaire au champ magnétique. Tout en haut à gauche on voit une raie simple (singulet) photographiée en l'absence de champ magnétique et immédiatement au-dessous l'image de son éclatement en un triplet

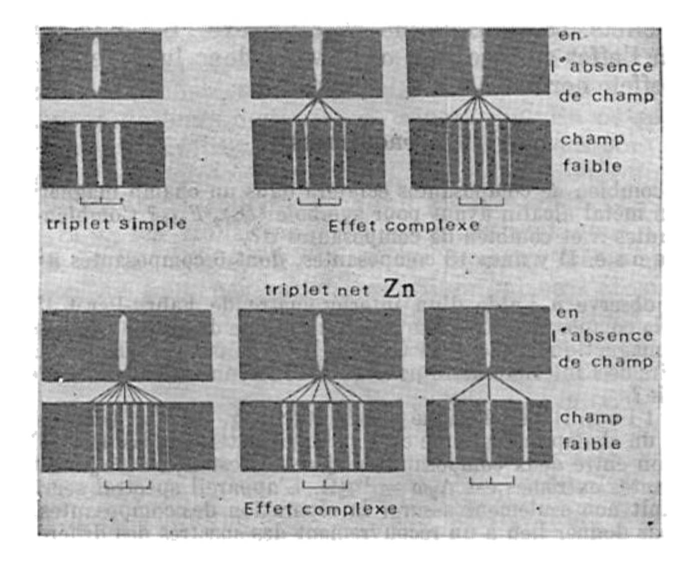


Fig. 74

de Lorentz dans un champ magnétique. En haut à droite on voit la photographie du doublet jaune du sodium et au-dessous son éclatement dans un champ magnétique faible (ce cas a fait l'objet du calcul donné dans le sous-paragraphe 4). Enfin en bas du dessin on a reproduit le triplet du zinc et au-dessous les éclatements de ses raies dans un champ magnétique faible.

8. Donnons une définition *quantitative* des champs magnétiques *faibles* et *forts*. Posons que $\Delta\omega$ désigne la largeur du multiplet considéré. Si $\Omega \ll \Delta\omega$, le champ sera dit faible et si $\Omega \gg \Delta\omega$, le champ sera dit fort. En passant aux longueurs d'onde on a $\Delta\omega = |2\pi c\Delta(1/\lambda)| = 2\pi c\Delta\lambda/\lambda^2$. En substituant à Ω sa valeur $\Omega = eB/2\mu c$, la condition de champ faible est

$$B \ll 4\pi\mu c^2\Delta\lambda/e\lambda^2. \tag{41.12}$$

Par exemple, pour la raie jaune de résonance du sodium, avec $\lambda = 593$ nm, $\Delta\lambda = 0{,}6$ nm la formule (41.12) donne pour les champs

faibles $B \ll 3{,}7\cdot 10^5$ Gs. On comprend pourquoi même dans un champ de l'ordre de 10^4 Gs la raie jaune donne un effet anormal. Pour la raie L_α de la série de Lyman de l'hydrogène $\lambda = 121{,}6$ nm, $\Delta\lambda = 5{,}3\cdot 10^{-4}$ nm et le champ sera faible si $B \ll 8000$ Gs. Pour la raie H_α de la série de Balmer $\lambda = 656$ nm, $\Delta\lambda = 0{,}0227$ nm (revoir problème du paragraphe précédent) et $B \ll 1{,}1\cdot 10^4$ Gs. Ces données montrent à quel point sont inadéquates les dénominations « effet normal » et « effet anormal » que l'on utilise encore pour qualifier les deux effets Zeeman simple et complexe. Le plus souvent on a affaire à l'effet complexe et on devrait donc lui réserver le qualificatif d'effet normal.

PROBLÈMES

1. En combien de composantes éclatera dans un champ magnétique faible la raie d'un métal alcalin ayant pour symbole $^2D_{5/2}{}^2F_{7/2}$? Combien y aura-t-il de composantes π et combien de composantes σ?

R é p o n s e. Il y aura 18 composantes, dont 6 composantes π et 12 composantes σ.

2. On observe à l'aide d'un interféromètre de Fabry-Pérot l'éclatement Zeeman de la composante de courte longueur d'onde de la raie D du sodium dans un champ magnétique $B = 5000$ Gs. Quelle doit être la distance d entre les plans miroirs de l'interféromètre pour qu'il convienne à l'étude de l'éclatement de cette raie?

S o l u t i o n. Il a été indiqué plus haut (cf. fig. 73) que la raie D du sodium éclate dans un champ magnétique en six composantes équidistantes; la distance de séparation entre deux composantes adjacentes est $\Delta_1\omega = {}^2/_3\Omega$ et celle entre les composantes extrêmes est $\Delta_2\omega = {}^{10}/_3\Omega$. L'appareil spectral servant à l'observation doit non seulement assurer la séparation des composantes, mais doit encore ne pas donner lieu à un recouvrement des spectres des différents ordres. Il s'ensuit que la région de dispersion de l'appareil ne doit pas être inférieure à $\Delta_2\omega$. La région de dispersion $\Delta\omega$ exprimée en fréquences est donnée par la relation $\Delta\omega = \omega/m$, où $m = 2d/\lambda = \omega d/\pi c$ est l'ordre du spectre. On doit donc avoir

$$\omega/m \geqslant \Delta_2\omega,$$

soit

$$m \leqslant \frac{\omega}{\Delta_2\omega}. \tag{41.13}$$

On trouve finalement

$$d \leqslant \frac{3}{5}\,\frac{\pi\mu c^2}{eB} = 0{,}64 \text{ cm}.$$

Le pouvoir de résolution de l'appareil est donné par l'expression $N_{eff}m$, où N_{eff} est le nombre effectif de faisceaux qui interfèrent. Pour assurer la résolution des composantes il faut que

$$N_{eff}\, m \geqslant \frac{\omega}{\Delta_1\omega}, \quad \text{i.e.} \quad N_{eff} m \geqslant 5\,\frac{\omega}{\Delta_2\omega}.$$

Avec (41.13) on trouve $N \geqslant 5$.

§ 42. Résonance magnétique

1. En présence d'un champ magnétique chaque niveau énergétique de l'atome éclate en $2J + 1$ sous-niveaux. Les sous-niveaux d'un seul et même niveau se distinguent les uns des autres par les nombres quantiques m_J caractérisant les projections du vecteur $\boldsymbol{J}$ sur la direction du champ magnétique. L'éclatement des raies spectrales que l'on observe dans l'effet Zeeman résulte de transitions quantiques entre les sous-niveaux provenant de niveaux éclatés différents. Les transitions spontanées entre les sous-niveaux d'un même niveau sont peu probables puisque leurs probabilités sont proportionnelles au cube de leurs distances de séparation. En outre, lorsqu'il n'y a qu'un seul électron de valence ces transitions sont interdites par la règle de sélection $\Delta L \neq 0$. Or les règles de sélection ne concernent que les transitions radiatives d'atomes *isolés*. Les transitions stimulées, i.e. les transitions qui se produisent sous l'action de champs de forces extérieurs, peuvent se produire même si les règles de sélection ne sont pas satisfaites. Si on place l'atome dans un champ de force adéquat, les probabilités des transitions quantiques considérées peuvent augmenter notablement. C'est le cas des transitions entre les sous-niveaux formés par éclatement dans un champ magnétique permanent $\boldsymbol{B}$ d'un seul et même niveau. Ces transitions s'effectuent avec des vitesses notables si on superpose au champ $\boldsymbol{B}$ un champ magnétique transversal variable de faible intensité. L'ensemble des phénomènes et des méthodes d'étude liés aux transitions stimulées de ce type sont désignés sous le nom de *résonance magnétique*.

2. Il est impossible de donner dans un cours de physique générale la théorie quantique rigoureusement conséquente de la résonance magnétique, car elle fait intervenir la question du calcul des probabilités des transitions stimulées des systèmes d'atomes d'un état à un autre. Mais pour donner une vue claire de la nature du phénomène il suffit de recourir à des considérations classiques simples. Cette approche est d'autant mieux justifiée que le phénomène de la résonance magnétique avait été prévu et expliqué en théorie classique. Les résultats des prévisions théoriques furent confirmés par l'expérience et c'est pour cela qu'on fait appel, encore aujourd'hui, aux modèles classiques pour donner une image concrète du phénomène. L'interprétation quantique conséquente fut élaborée plus tard.

Considérons une particule de moment cinétique $\boldsymbol{J}$ et de moment magnétique $\vec{\mathfrak{m}}$ se trouvant dans un champ magnétique $\boldsymbol{B}$, permanent et homogène. Posons que ces deux moments soient liés par la relation $\vec{\mathfrak{m}} = g\boldsymbol{J}$, g étant le *rapport gyromagnétique*. Cela étant on suppose que l'unité de $\boldsymbol{J}$ est la constante de Planck $\hbar$ et que l'unité du mo-

ment magnétique est le magnéton de Bohr $e\hbar/(2\mu_e c) = 9{,}273 \times \times 10^{-21}$ erg/Gs ou le magnéton nucléaire $e\hbar/(2\mu_p c) = 5{,}050 \times \times 10^{-24}$ erg/Gs, suivant que le moment magnétique de la particule est déterminé par les électrons ou par les noyaux atomiques. Ici μ_e et μ_p dénotent respectivement les masses de l'électron et du proton. C'est pour cela que le magnéton nucléaire est 1836 fois plus petit que le magnéton électronique de Bohr. Dans le cas où on aurait à exprimer le rapport gyromagnétique en unités absolues on le notera g_{abs} et on écrira donc $\vec{\mathfrak{m}} = g_{abs}J$. La quantité g est sans dimension, tandis que g_{abs} a les dimensions de $e/\mu c$. Par exemple, pour le spin électronique $g_{abs} = e/\mu_e c$, tandis que $g=2$. Les spins des noyaux sont du même ordre de grandeur que ceux des atomes et s'expriment par des nombres entiers ou demi-entiers, tandis que les moments magnétiques nucléaires sont des milliers de fois plus petits que le moment magnétique de l'électron. Pour des particules d'espèces différentes les valeurs de g peuvent être différentes, mais cette circonstance importe peu pour nos besoins.

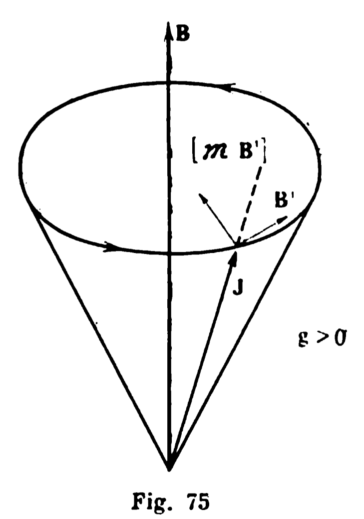


Fig. 75

Une particule placée dans un champ magnétique est soumise à un moment de rotation $[\vec{\mathfrak{m}}B]$, de sorte que

$$\dot{J} = [\vec{\mathfrak{m}}B] = g_{abs}[JB].$$

C'est l'*équation de la toupie*. En régime établi le moment J et avec lui le moment $\vec{\mathfrak{m}}$ exécuteront une précession régulière stimulée (fig. 75) avec une vitesse angulaire

$$\Omega = -g_{abs}B. \qquad (42.1)$$

Superposons maintenant au champ B un champ magnétique faible B' orthogonal à B et tournant autour de ce dernier. La particule en précession sera alors soumise à un moment de force supplémentaire variable $[\vec{\mathfrak{m}}B']$. Suivant son orientation ce moment fera varier l'angle entre les vecteurs J et B. Si la vitesse de précession Ω et la vitesse angulaire de rotation Ω' du champ B' sont notablement différentes, ces deux rotations seront constamment déphasées et le moment de rotation $[\vec{\mathfrak{m}}B']$ subira de petites variations périodiques, rapidement variables en module et en direction. Ces variations ne pro-

voqueront que de petites variations de l'angle d'inclinaison entre J et B variant rapidement et périodiquement dans le temps. Elles tantôt rapprocheront et tantôt éloigneront en direction le vecteur J du vecteur B, de sorte qu'en moyenne la direction du vecteur J restera constante.

La situation est complètement différente lorsque Ω' coïncide avec Ω. C'est alors que se manifeste le phénomène de résonance magnétique. Ce phénomène consiste en ce que le moment de force supplémentaire s'exerce tout le temps dans le même sens et oriente le vecteur J soit parallèlement, soit antiparallèlement au champ magnétique permanent B. La longueur du vecteur J reste constante, mais sa projection J_z sur la direction du champ B varie.

3. Les résultats de l'étude quantique ne diffèrent de ceux de l'étude classique qu'en ce que dans le cas quantique la projection J_z est quantifiée, ce qui entraîne des variations temporelles par sauts, tandis que dans le cas classique ces variations sont continues. L'étude classique conduit à conclure que la résonance ne peut s'observer qu'à la seule fréquence Ω. En appliquant le principe de correspondance, on est amené à la règle de sélection $\Delta m_J = \pm 1$ qui doit être vérifiée dans les transitions quantiques entre les sous-niveaux formés par éclatement d'un niveau par effet Zeeman. Cela signifie que seules les transitions entre sous-niveaux voisins sont permises. En effet, la distance entre les sous-niveaux est égale à $\Delta\mathscr{E} = g B \Delta m_J \hbar$, de sorte que la fréquence de la transition sera $\omega = \Delta\mathscr{E}/\hbar = g B \Delta m_J$, résultat qui coïncidera avec $\pm\Omega$, à condition que $\Delta m_J = \pm 1$. Il est bien évident que cette règle de sélection est une conséquence de la loi de la conservation du moment cinétique, puisqu'on peut assimiler la transition stimulée entre des sous-niveaux à l'interaction d'un système atomique avec un photon de fréquence Ω (cf. § 39).

4. Pour observer commodément la résonance magnétique il est préférable d'utiliser à la place du champ tournant B' un champ variant sinusoïdalement dans une même direction orthogonale au champ principal B. En principe rien ne change puisqu'on représente le champ B' sous forme d'une superposition de deux champs de même intensité tournant avec la même vitesse angulaire Ω mais dans des sens opposés. Le champ tournant à l'encontre du sens de la précession de la particule, exerce sur celle-ci une action oscillante rapide qui ne joue aucun rôle dans le phénomène considéré. Seul importe le champ qui fait varier toujours dans le même sens l'angle entre J et B, i.e. le champ qui tourne dans le sens de la précession du vecteur J. La fréquence circulaire Ω requise pour faire apparaître la résonance magnétique est donnée par la formule (42.1). En passant aux longueurs d'onde, on en tire

$$\lambda = \frac{2\pi c}{\Omega} = \frac{2\pi c}{|g_{abs}|\, B}. \qquad (42.2)$$

Dans le cas où les moments magnétique et mécanique de la particule sont déterminés par les électrons de l'enveloppe électronique de l'atome la résonance magnétique porte le nom de *résonance paramagnétique électronique* et s'ils sont déterminés par les noyaux atomiques on l'appelle *résonance magnétique nucléaire*. On connaît encore les résonances *ferromagnétique*, *antiferromagnétique* et *diamagnétique* (*cyclotronique*) dont il sera question plus loin. La résonance magné-

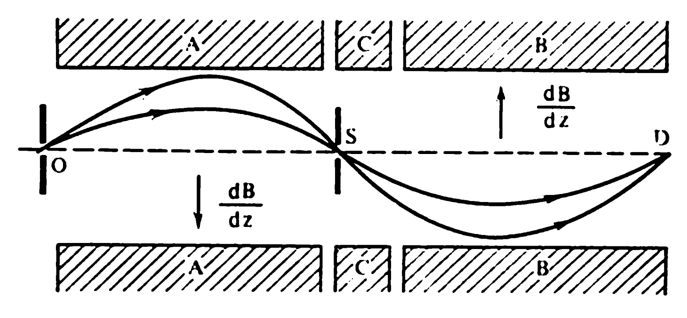


Fig. 76

tique est largement utilisée pour la détermination des moments magnétiques des atomes et des noyaux atomiques, pour l'étude de la structure des molécules et des cristaux, etc.

Le spin des électrons étant égal à 1/2, $g_{abs} = e/\mu c$. Si l'intensité du champ magnétique $B \approx 3 \cdot 10^3$ Gs, la formule (42.2) donne dans ce cas

$$\lambda = \frac{2\pi\mu c^2}{eB} \approx 3{,}5 \text{ cm}.$$

La fréquence correspondante est

$$\nu = \frac{c}{\lambda} \approx 10^4 \text{ MHz}.$$

Ce sont des fréquences de la gamme des microondes ($\nu > 30$ MHz, $\lambda < 1$ m) appelées hyperfréquences. La résonance magnétique nucléaire doit s'observer, dans des champs magnétiques de même intensité, à des fréquences beaucoup plus basses de l'ordre de quelques mégahertz (longueurs d'onde des milliers de fois plus grandes).

5. Dans l'ordre historique, les premières observations de la résonance magnétique nucléaire avec des faisceaux atomiques ou moléculaires neutres ont été faites en 1938 par la méthode élaborée par Rabi (né en 1898). Les moments mécanique et magnétique des enveloppes électroniques des atomes ou des molécules composant le faisceau doivent être compensés, afin que les deux moments soient purement nucléaires. Dans la méthode de Rabi les atomes ou les molécules du faisceau après avoir traversé une fente (fig. 76) pénètrent d'abord dans le champ inhomogène d'un aimant A à gradient per-

pendiculaire à la fente et dirigé, par exemple, vers le bas. A la suite de l'aimant A se trouve l'aimant C qui crée un fort champ homogène et permanent B dans lequel se trouve une fente S. On trouve ensuite un troisième aimant B, en tout point analogue à l'aimant A, dont le champ inhomogène a un gradient orienté en sens inverse du gradient du champ créé par l'aimant A, donc orienté vers le haut. La construction des aimants A et B est analogue à celle des aimants que l'on utilise dans les expériences de Stern et Gerlach (cf. § 36).

Si les aimants A et B n'existaient pas, le faisceau de particules dirigé suivant l'axe de l'appareil, après avoir traversé la fente S tomberait sur le détecteur D. Mais dans des champs inhomogènes, du fait que les particules possèdent des moments magnétiques, elles sont soumises à l'action de forces transversales $F = \mathfrak{m}_z(dB/dz)$ orientées parallèlement ou antiparallèlement au gradient du champ magnétique suivant le signe de la projection $\mathfrak{m}_z$ sur la direction du champ. Ces forces incurvent les trajectoires des particules. La source envoie dans l'appareil des particules ayant des vitesses et des directions différentes. Seules les particules qui arrivent à passer à travers la fente S présentent de l'intérêt. La force F qui agit sur une particule parcourant sa trajectoire ne change pas, car la projection $\mathfrak{m}_z$ ne varie pas. En conséquence, on peut poser, en première approximation, que les trajectoires des particules, dans le champ de l'aimant A, sont des cercles dont le rayon R est donné par l'équation $v^2/R = f$, où f est la force normale qui s'exerce sur l'unité de masse de la particule dans le champ inhomogène de l'aimant A. Toutes les particules ainsi orientées qui ne sont pas arrêtées par les bords du diaphragme passent au travers de la fente S.

Dans le champ de l'aimant B, les projections $\mathfrak{m}_z$ n'ayant pas changé, les trajectoires des particules seront à nouveau des cercles mais infléchis dans le sens opposé. Par un choix convenable de la configuration du champ de l'aimant B, les déviations des particules, occasionnées par l'aimant A, seront compensées par l'aimant B et les particules parviendront jusqu'au détecteur D. La figure 76 représente deux trajectoires de ce type. Dans ces conditions le détecteur D enregistrera le maximum du flux de particules.

On a supposé jusqu'ici que tous les champs magnétiques étaient permanents. Superposons maintenant au champ permanent fort B produit par l'aimant C un faible champ magnétique transversal B' de fréquence radio, qui varie harmoniquement dans le temps avec la fréquence ω. Ce champ stimulera les transitions quantiques des particules; du fait de ces transitions les projections $\mathfrak{m}_z$ de certaines particules seront modifiées, ce qui modifiera les forces s'exerçant sur ces particules dans le champ inhomogène de l'aimant B. Dans la région de l'espace occupée par le champ de l'aimant B les particules s'écarteront de leurs trajectoires et ne parviendront plus jusqu'au détecteur D.

Supposons d'abord que $\omega < \Omega$. A mesure que ω augmente, l'intensité N du faisceau de particules tombant sur le détecteur D diminuera. Pour $\omega = \Omega$ survient la résonance entre les oscillations du champ B' et la précession de Larmor des particules autour de la direction du champ permanent fort B. A ce moment l'intensité N du faisceau de particules tombant sur le détecteur atteint sa valeur minimale. Lorsqu'on dépasse la fréquence de résonance Ω et que la fréquence ω continue à augmenter, l'intensité N se met à augmenter.

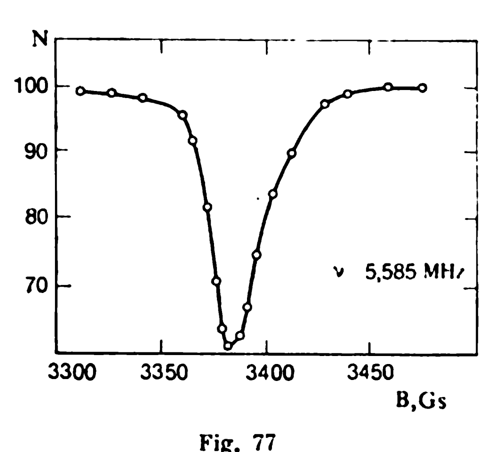


Fig. 77

Il résulte de la formule (42.1) que la résonance s'observe pour

$$\omega = |g_{abs}|\, B = g\,\frac{eB}{2\mu c} \qquad (42.3)$$

ou

$$\nu = g\,\frac{eB}{4\pi\mu c}, \qquad (42.4)$$

où on entend par μ la masse du proton s'il s'agit de résonance magnétique nucléaire ou la masse de l'électron s'il s'agit de la résonance paramagnétique électronique. Ayant déterminé la fréquence ν et l'intensité du champ B au minimum, la formule (42.4) permet de calculer le rapport gyromagnétique g pour les noyaux atomiques. La précision relative de ces mesures atteint 0,001 %. Connaissant le spin du noyau, on peut calculer son moment magnétique $\mathfrak{m}$. On reviendra sur cette question en Physique nucléaire.

Dans les expériences il est plus commode de maintenir constante la fréquence du champ magnétique B' et de faire varier lentement l'intensité du champ fort B de part et d'autre d'une certaine valeur moyenne. Pour y arriver on procède à la modulation du champ de l'aimant C avec une basse fréquence (50 Hz) produite par des bobines alimentées par le réseau urbain. La figure 77 illustre l'acuité du minimum de résonance qui fut obtenu dans une étude concernant les noyaux ^{7}Li où on a utilisé un faisceau de molécules à spins électroniques compensés. En abscisses on a porté l'intensité en gauss du champ « constant » B et en ordonnées on a porté l'intensité relative N du faisceau de particules enregistrées par le détecteur D. La fréquence du champ de radiofréquence à laquelle fut observé le minimum de résonance est $\nu = 5{,}585$ MHz.

6. L'une des principales applications de la résonance magnétique fondée sur l'utilisation des faisceaux atomiques est la mesure du moment magnétique $\mathfrak{m}_{el}$ de l'électron. On s'aperçoit que cette quantité ne coïncide pas exactement avec le magnéton de Bohr. En

écrivant le moment magnétique de l'électron sous la forme $\mathfrak{m}_{el} = \frac{1}{2}\, g\mathfrak{m}_B$, on constate que le facteur $g/2$ n'est pas exactement égal à 1, comme l'implique la théorie de Dirac. Les mesures effectuées sur des jets d'atomes d'hydrogène se trouvant à l'état fondamental ont fourni le résultat suivant :

$$\frac{1}{2}\, g_{exp} = 1{,}0011596524 \pm 20\cdot 10^{-11}$$

qui est en bon accord avec la valeur théorique calculée à l'aide d'ordinateurs

$$\frac{1}{2}\, g_{théor} = 1{,}00115965238 \pm 26\cdot 10^{-11}.$$

C'est à cette valeur que conduit l'électrodynamique quantique contemporaine. Par conséquent, selon l'électrodynamique quantique le moment magnétique de l'électron est égal à

$$\mathfrak{m} = 1{,}0011596522091 \pm 31\cdot 10^{-11}\ \mathfrak{m}_B.$$

Cette valeur est appelée *moment magnétique anomal* de l'électron.

La méthode des jets moléculaires ne s'applique qu'aux particules *neutres*, ce qui restreint son domaine d'application. En effet, les particules portant la charge e et se déplaçant avec la vitesse v seraient soumises à la force de Lorentz $e/c\ [vB]$ qui déterminerait un déplacement latéral, variable suivant la vitesse des particules.

La méthode de résonance magnétique utilisant des jets de molécules neutres présente l'avantage d'être très précise. Un autre avantage est que le champ de radiofréquence exerce son action sur des particules *libres*. Mais la technique expérimentale est malaisée et exige la mise en œuvre de techniques du vide sophistiquées. C'est pour cela qu'on n'utilise que rarement la méthode des jets moléculaires.

7. C'est pour ces raisons que l'on utilise largement une autre méthode de résonance magnétique de technique expérimentale beaucoup plus simple, dans laquelle on utilise des *quantités macroscopiques* de substance à l'état solide, liquide ou gazeux. Les atomes, les molécules ou les ions de la substance doivent posséder des moments magnétiques électroniques ou nucléaires. Lorsqu'on place l'échantillon étudié dans un champ magnétique permanent de grande intensité B, par suite des chocs interatomiques il s'établit pour l'échantillon, au bout d'un temps court, un état d'équilibre caractérisé par l'aimantation de la substance. C'est le phénomène de *paramagnétisme électronique* ou *nucléaire* et c'est pour cela que la méthode est appelée résonance paramagnétique électronique ou résonance magnétique nucléaire.

Une particule dont le moment magnétique est orienté suivant le champ possède une énergie moindre qu'une particule identique dont le moment magnétique est antiparallèle au champ. Selon la formule de Boltzmann, à l'état d'équilibre, le nombre de particules

orientées parallèlement au champ doit être supérieur au nombre de particules orientées antiparallèlement au champ. Cela signifie que les sous-niveaux inférieurs formés par effet Zeeman auront une population plus grande que les niveaux supérieurs.

Evaluons la différence des populations des sous-niveaux à la température ordinaire $T = 293$ K en supposant que le moment magnétique d'une particule est égal à un magnéton de Bohr $\mathfrak{m} = 9{,}27 \times \times 10^{-21}$ erg/Gs et que $B = 5 \cdot 10^3$ Gs. En notant n_2 et n_1 les nombres de particules occupant respectivement le sous-niveau supérieur $\mathscr{E}_2$ et le sous-niveau inférieur $\mathscr{E}_1$, l'application de la formule de Boltzmann fournit le résultat suivant:

$$\frac{n_1}{n_2} = e^{\frac{\mathscr{E}_2 - \mathscr{E}_1}{kT}} \approx 1 + \frac{\mathscr{E}_2 - \mathscr{E}_1}{kT} = 1 + \frac{2\mathfrak{m}B}{kT},$$

de sorte que

$$\frac{n_1 - n_2}{n} = \frac{2\mathfrak{m}B}{kT} = 2{,}3 \cdot 10^{-3}.$$

Ainsi, la différence des populations de deux sous-niveaux voisins ne représente que 0,2 % environ de la population de l'un de ces sous-niveaux.

Malgré la petitesse de cette différence, elle peut se manifester à l'échelle macroscopique, vu que le nombre de particules occupant chacun des sous-niveaux est très grand. Lors de la transition stimulée d'une particule du sous-niveau supérieur au sous-niveau inférieur il se produit l'émission d'un quantum d'énergie correspondant à la différence des énergies des sous-niveaux considérés. Les transitions allant du sous-niveau inférieur au sous-niveau supérieur empruntent l'énergie nécessaire au champ de radiofréquence. Les transitions spontanées radiatives étant relativement peu nombreuses, on n'en tiendra pas compte. En définitive l'absorption d'énergie prévaudra sur l'émission, quoique les probabilités des transitions directes et inverses soient égales (cf. t. IV, § 119). Cette différence sera maximale lorsque la fréquence du champ de radiofréquence coïncidera avec la fréquence de Larmor de la particule en précession, i.e. à la résonance. La résonance magnétique apparaît lorsque l'énergie du champ de radiofréquence absorbée par les particules, à proximité de la fréquence de Larmor, sera minimale.

Par suite de la prédominance des transitions des niveaux inférieurs aux niveaux supérieurs les énergies des sous-niveaux tendent à s'égaliser, mais les processus de relaxation, qui cherchent à rétablir l'état d'équilibre initial, s'y opposent. Pour que la résonance magnétique puisse se manifester avec une acuité suffisante il faut que la période des oscillations HF soit grande devant le temps de relaxation au cours duquel se rétablit l'état d'équilibre.

8. La résonance paramagnétique électronique fut découverte en 1944 par E. C. Zavoïsky (1907-1976). Il obtint ses premiers résultats en travaillant avec des sels du groupe du fer. Par la suite la gamme des substances étudiées s'étendit grandement. Zavoïsky travaillait avec des ondes radio de la gamme décimétrique, ce qui permettait, conformément à la formule (42.2), d'utiliser des champs magnétiques B de faible intensité. Avec le développement de la technique des ondes courtes on commence à utiliser des ondes centimétriques.

Dans les radiospectroscopes modernes on fixe la fréquence du signal radio et on module le champ magnétique B avec une fréquence basse (50 Hz). Un schéma de l'appareil est représenté sur la figure 78. L'électroaimant NS est alimenté par un courant continu et produit un fort champ magnétique permanent, que l'on module à l'aide des bobines KK alimentées par un courant alternatif d'une fréquence de 50 Hz. L'échantillon A ayant un volume de quelques mm³ est placé dans une cavité résonante R accordée à une longueur d'onde $\lambda \sim 3$ cm. Les ondes électromagnétiques de cette longueur sont générées par un klystron réflex et sont introduites dans la cavité résonante R à travers le guide d'ondes F. Après avoir été partiellement absorbées par l'échantillon A, ces ondes sont dirigées à travers un second guide d'ondes vers un détecteur cristallin à silicium-tungstène D ; après détection ces ondes peuvent être amplifiées. Pour déceler l'absorption paramagnétique des ondes on réunit le détecteur à un galvanomètre sensible. Une sensibilité notablement plus drande peut être atteinte en explorant le signal détecté et amplifié par l'amplificateur Y sur l'écran d'un oscillographe (qui n'est pas représenté sur la figure 78). Le signal est appliqué aux plaques de déflexion verticale de l'oscillographe ; la déflexion horizontale fixe les valeurs instantanées du champ magnétique B ; cette dernière déflexion doit être accordée en phase avec la modulation du champ magnétique B. On observe sur l'écran de l'oscillographe une courbe d'absorption paramagnétique analogue à celle représentée sur la figure 77. Les radiospectroscopes modernes permettent de déceler la

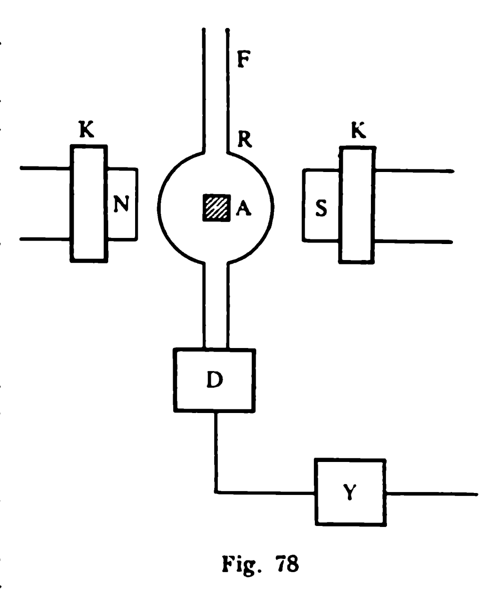


Fig. 78

résonance paramagnétique électronique à la température ambiante dans des échantillons contenant jusqu'à 10^{11}-10^{12} particules paramagnétiques.

9. La résonance magnétique nucléaire avec l'absorption d'ondes radio courtes par des quantités macroscopiques des substances fut observée pour la première fois en 1945 par Purcell (né en 1912), Torrey et Pound (né en 1919), et indépendamment par Bloch (né en 1905), Hansen (1909-1949) et Packard. Quoique la résonance magnétique nucléaire ne diffère pas, en principe, de la résonance électronique, vu la nécessité d'utiliser dans les expériences un champ électromagnétique HF, ayant des longueurs d'onde très différentes, la méthode expérimentale présente des particularités. Le schéma de la figure 79 illustre le principe de l'un des montages expérimentaux. La principale particularité de la méthode est que l'échantillon *A* est placé dans une bobine branchée en série avec la bobine du circuit oscillant d'un générateur haute fréquence. L'électroaimant NS alimenté en courant continu crée un champ magnétique permanent de grande intensité. Pour créer le champ magnétique orthogonal de modulation, on utilise une bobine alimentée en courant alternatif d'une fréquence de 50 Hz. A la résonance l'absorption du champ électromagnétique HF s'accentue, ce qui détermine une diminution notable du facteur de qualité du circuit oscillant du générateur allant jusqu'au blocage de la génération. Le signal HF qui apparaît dans le circuit extérieur, couplé par inductance au circuit oscillant du générateur, est alors détecté et amplifié par l'amplificateur *Y*. Ensuite on l'explore sur l'écran d'un oscillographe exactement comme on le fait dans le cas de la résonance paramagnétique électronique.

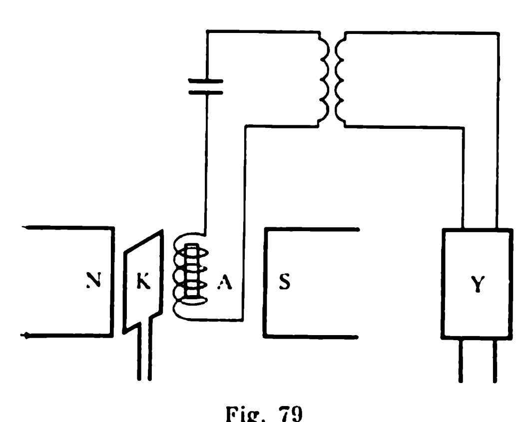


Fig. 79

10. Notons, pour conclure, qu'outre les deux types de résonance décrits ci-dessus on utilise aussi la *résonance ferromagnétique* liée aux changements d'orientation des moments magnétiques électroniques dans les domaines ou entre les domaines des corps ferromagnétiques et la *résonance antiferromagnétique* liée aux changements d'orientation des moments magnétiques de spin dans la substance antiferromagnétique. Un cas particulier de ce dernier type de résonance est la *résonance ferrimagnétique* qui se manifeste dans les substances ferrimagnétiques qui sont des antiferromagnétiques dont les moments

magnétiques antiparallèles des spins du réseau ne sont pas exactement compensés, ce qui conduit à l'apparition des propriétés ferromagnétiques. Signalons encore qu'on associe formellement à la résonance magnétique la *résonance diamagnétique* (*cyclotronique*) ayant cependant une nature physique totalement différente.

§ 43. L'effet Stark

1. L'effet Stark (1874-1957) consiste en ce que l'application d'un champ électrique détermine le déplacement et le dédoublement en sous-niveaux des niveaux énergétiques des atomes, des molécules et des cristaux. Cela se manifeste dans le dédoublement et le déplacement des raies spectrales dans les spectres d'émission et d'absorption des substances indiquées. On a déjà mentionné l'existence de cet effet dans le t. IV, § 93, en signalant les difficultés expérimentales que Stark eut à surmonter pour observer l'effet. Stark a découvert l'effet qui porte son nom et l'étudia en détail sur les raies spectrales de la série de Balmer de l'hydrogène. Par la suite l'effet Stark a été observé avec d'autres atomes.

On se rendit compte dès le début que la théorie classique ne pouvait expliquer l'effet Stark. Une théorie de l'effet Stark fondée sur la théorie semi-classique de Bohr fut élaborée en 1916 indépendamment par K. Schwarzschild (1874-1916) et par P.S. Epstein (1886-1966). Les principaux résultats de cette théorie ont été confirmés par une théorie quantique conséquente élaborée en 1926 par Schrödinger. Dans les deux théories on met en œuvre les méthodes de calcul de la théorie des perturbations qui ont été initialement développées pour la mécanique céleste par Lagrange (1736-1813), Laplace (1749-1827) et d'autres savants, puis adaptées aux besoins de la mécanique quantique. Les calculs sont assez laborieux et ne peuvent être reproduits ici. On se contentera d'exposer quelques considérations qualitatives ainsi que les résultats définitifs de la théorie. On ne considérera que l'effet Stark se manifestant dans les atomes en supposant que le champ électrique E est uniforme.

2. A l'aide de considérations classiques simples on arrive à comprendre quelles doivent être les polarisations des composantes résultant de l'éclatement des raies spectrales lorsqu'on soumet la source lumineuse à l'action d'un champ électrique extérieur E. Dans le champ électrique, la fréquence des vibrations de la source lumineuse élémentaire (l'électron) dépend de ce que les vibrations sont parallèles ou perpendiculaires au champ E. Dans tous les cas, vu la transversalité des vibrations lumineuses, on ne pourra observer que des vibrations perpendiculaires à la ligne d'observation. Si cette dernière est orthogonale au champ E, les vibrations satisfaisant à cette condition peuvent être soit parallèles, soit perpendiculaires au champ E. Comme les fréquences de ces vibrations sont généralement différentes, toutes les raies du spectre observé seront polarisées rectilignement : certaines raies seront polarisées parallèlement au champ E (composantes π) et les autres perpendiculairement à ce champ (composantes σ).

Dans le cas où la ligne d'observation est parallèle au champ E, toutes les vibrations donnant lieu à une émission de lumière seront perpendiculaires au champ E, et on n'observera dans le spectre que des composantes σ; aucune de ces composantes n'est polarisée puisque la force qu'exerce le champ électrique E sur un électron oscillant ne dépend ni de la grandeur ni de la direction de sa vitesse de déplacement. C'est là une différence essentielle entre le champ électrique et le champ magnétique. La force qu'exerce un champ magnétique sur l'électron est proportionnelle à sa vitesse v et change de sens si la vitesse v de l'électron change de sens. C'est pour cela que cette force peut faire varier les vitesses angulaires des rotations circulaires de l'électron, en lesquelles on peut décomposer son mouvement vibratoire. Ces variations des vitesses angulaires dépendent du sens de rotation de l'électron et c'est à cela qu'est lié l'effet Zeeman longitudinal. Comme le champ électrique ne peut provoquer de telles variations,

les composantes formées par effet Stark *ne sont pas polarisées* lors d'observations longitudinales. Si on les observe suivant une direction oblique par rapport au champ E, elles seront *partiellement* polarisées.

3. Il existe deux variétés d'effet Stark, suivant que l'atome possède ou ne possède pas de moment électrique dipolaire p (en l'absence de champ électrique E). Dans le premier cas l'application du champ électrique E communique à l'atome une énergie supplémentaire (pE), qui sera proportionnelle à l'intensité du champ E, si on se limite aux termes linéaires par rapport à E. Le déplacement et le dédoublement des raies spectrales sont alors proportionnels au champ électrique. C'est cet effet qu'observa Stark.

Dans le second cas l'atome n'a pas de moment électrique propre, et lorsqu'on le soumet à l'action d'un champ électrique, celui-ci fait apparaître chez l'atome un *moment dipolaire induit* $p = \beta E$, où β est la polarisabilité de l'atome que l'on peut calculer par les méthodes de la mécanique quantique. Lorsqu'on fait croître le champ électrique de 0 à E le moment dipolaire de l'atome augmente aussi de 0 jusqu'à p. L'atome est soumis alors à un travail $^1/_2\,(pE) = {}^1/_2\beta E^2$ qui sert à augmenter l'énergie potentielle de l'atome dans le champ électrique. (L'apparition du facteur 1/2 est déterminée par la même raison que dans le cas analogue de calcul de l'énergie potentielle d'un corps déformé de façon élastique conformément à la loi de Hooke.) Il s'ensuit que le déplacement et le dédoublement des raies spectrales seront proportionnels à E^2. L'effet Stark correspondant est dit *quadratique*. Comme l'effet quadratique est de toute évidence beaucoup plus faible que l'effet linéaire, il ne fut décelé que plus tard.

Il est évident qu'un atome possédant un moment dipolaire propre acquiert un moment dipolaire supplémentaire lorsqu'on le soumet à l'action d'un champ électrique. En première approximation on posera que ce moment supplémentaire est proportionnel à l'intensité du champ E. Dans ces conditions il se produit une superposition des effets Stark *linéaire* et *quadratique*. La figure de dédoublement des niveaux est alors asymétrique: tous les sous-niveaux sont déplacés du côté des énergies décroissantes, les déplacements étant d'autant plus importants qu'ils se trouvent plus haut placés dans l'échelle des énergies. Les raies spectrales sont déplacées vers le rouge, mais les déplacements sont petits. Par exemple, le déplacement de l'une des composantes de Stark de la raie H_α est d'environ 1 cm^{-1}, tandis que la distance de séparation des composantes extrêmes de Stark de cette raie est égale à 200 cm^{-1} *).

Pour l'hydrogène si le champ appliqué n'excède pas 10^5 V/cm, on peut négliger complètement l'effet Stark quadratique. Dans l'hydrogène le terme quadratique $\sim E^2$ ne commence à se manifester que dans des champs plus forts. Lorsque l'intensité du champ appliqué devient supérieure à $4\cdot 10^5$ V/cm environ, on voit apparaître un terme $\sim E^3$, qui a été calculé en même temps que le terme $\sim E^2$. Compte tenu de ces termes la théorie est en bon accord avec les résultats expérimentaux obtenus dans les champs les plus intenses que l'on ait réussi à produire, i.e. de l'ordre de 10^6 V/cm.

4. Le fait que dans l'hydrogène, ses isotopes (deutérium et tritium) et les ions hydrogénoïdes, l'effet Stark est *linéaire* tient à ce que le champ électrique créé par le noyau dans lequel se meut l'électron est un champ *coulombien*. Or dans un champ coulombien les niveaux énergétiques de l'électron sont dégénérés par rapport à l. Dans ce cas tous les états d'un atome monoélectronique ayant une même valeur du nombre quantique principal n, mais des valeurs de l différentes possèdent tous la *même énergie*. Les états dont la superposition reproduit un état quelconque avec un n donné, possèdent, même *dans un champ électrique extérieur nul*, des *moments électriques dipolaires propres*. Lorsqu'on applique un champ électrique extérieur la dégénérescence est partiellement levée et les niveaux énergétiques correspondant à des états différents sont inégalement déplacés. Mais tous ces déplacements et le dédoublement des raies spectrales qui

*) D'après la formule $1/\lambda = \nu/c$, 1 cm^{-1} = c Hz = $3\cdot 10^{10}$ Hz = $3\cdot 10^4$ MHz.

y est lié sont proportionnels au champ E et par suite l'effet Stark est *linéaire*.

Dans le cas d'atomes et d'ions plus compliqués ayant un seul électron de valence, on peut assimiler l'atome à un système monoélectronique. Mais dans ce cas le champ que crée le noyau et dans lequel se meut l'électron est déformé par les couches électroniques internes et de ce fait n'est plus coulombien. Dans un tel champ la dégénérescence par rapport à l ne se manifeste plus. Une étude plus poussée montre que dans chacun des états caractérisés par les nombres quantiques n et l le moment électrique moyen propre de l'atome est nul. Il s'ensuit que lorsqu'on applique un champ électrique le dédoublement des niveaux commence à partir des termes quadratiques en E. L'effet Stark observé est *quadratique*.

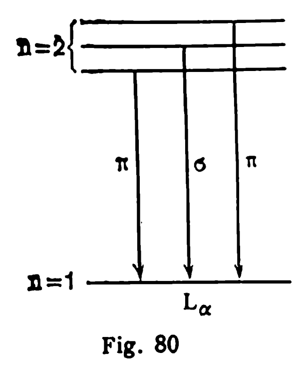


Fig. 80

5. Passons à une étude de l'effet Stark dans l'hydrogène. Si on convient de négliger le spin de l'électron, on pourra négliger l'interaction spin-orbite. Dans cette approximation le problème consiste à trouver une solution de l'équation de Schrödinger en tenant compte de l'énergie potentielle de l'atome dans un champ électrique extérieur. Le problème possède alors une *symétrie cylindrique*, l'axe de symétrie étant parallèle au champ électrique. Les coordonnées sphériques r, ϑ, φ conviennent bien pour traiter les problèmes caractérisés par une symétrie sphérique, mais sont d'un emploi peu commode dans le cas d'une symétrie cylindrique. Dans ce dernier cas il est préférable d'utiliser des coordonnées dites *paraboliques* qui possèdent la symétrie requise. La résolution de l'équation de Schrödinger pour l'atome d'hydrogène conduit à conclure que dans un champ électrique continu un niveau énergétique à nombre quantique principal n éclate en $2n - 1$ sous-niveaux. Les transitions entre ces sous-niveaux satisfaisant aux règles de sélection, déterminent les composantes qui apparaissent par suite de l'éclatement des raies spectrales de l'hydrogène sous l'action d'un champ électrique extérieur.

En présence d'un champ électrique extérieur la loi de la conservation du moment cinétique n'est pas vérifiée en général. Mais dans un champ électrique continu et uniforme doit être assurée la conservation de la *projection du moment cinétique* sur la direction du champ électrique. C'est pour cela que restent en vigueur les règles de sélection relatives au nombre quantique magnétique m_l, déterminant cette projection (comme convenu plus haut, on ne tient pas compte du spin s). Lorsque $\Delta m_l = 0$ on voit apparaître une composante π et lorsque $\Delta m_l = \pm 1$ c'est la composante σ qui fait son apparition. Ce sont ces règles de sélection qui définissent les transitions possibles.

C'est le schéma d'éclatement des raies d'hydrogène de la série de Lyman qui est le plus simple de tous. Les raies de cette série correspondent aux transitions des différents niveaux supérieurs sur le niveau $n = 1$ qui n'éclate pas ($2n - 1 = 1$). Le niveau $n = 2$ éclate en $2n - 1 = 3$ sous-niveaux. Les transitions de ces sous-niveaux sur le niveau $n = 1$ font apparaître trois composantes résultant de l'éclatement de la raie L_α de l'hydrogène. Ces transitions sont représentées sur la figure 80. Les niveaux $n = 3$ et $n = 4$ éclatent respectivement en $2n - 1 = 5$ et $2n - 1 = 7$ sous-niveaux. Les transitions de ces sous-niveaux au niveau $n = 1$ déterminent les composantes des raies L_β et L_γ. Le schéma de ces éclatements est représenté sur la figure 81 où les composantes π sont représentées par des traits gras dirigés vers le haut et les composantes σ par des traits semblables mais dirigés vers le bas. Les longueurs de ces traits indiquent les intensités relatives des composantes spectrales qui apparaissent lorsqu'on applique un champ électrique. On notera que dans le cas de la raie L_β la composante centrale manque, de sorte que cette raie éclate en quatre composantes seulement.

La raie L_γ éclate en 7 composantes, dont quatre sont du type π et trois du type σ. Ces résultats théoriques sont corroborés par les données de l'expérience, qui nécessite la mise en œuvre d'un appareil spectral à vide (travail dans l'ultraviolet !).

Le mode d'éclatement des raies spectrales de la série de Balmer de l'hydrogène est un peu plus compliqué car les transitions aboutissent aux trois sous-

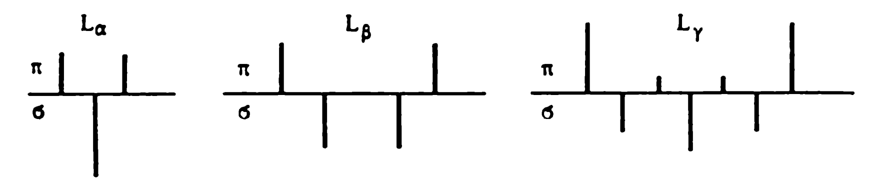


Fig. 81

niveaux formés par éclatement du niveau $n = 2$. Le niveau le plus proche $n = 3$ éclate en cinq sous-niveaux. Il s'ensuit que la raie de Balmer H_α déterminée par les transitions du niveau $n = 3$ au niveau $n = 2$ éclate en 15 composantes, comme le montre la figure 82. La raie H_β éclate en 20 composantes, la raie H_γ en 27, la raie H_δ en 32 composantes, etc. (Lors de l'éclatement des raies

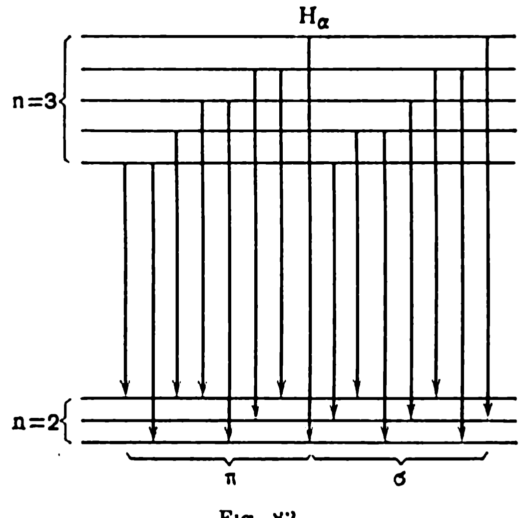


Fig. 82

H_β et H_δ les composantes centrales manquent, ce qui réduit le nombre des composantes d'une unité: 20 au lieu de 21 et 32 au lieu de 33.) L'éclatement des raies H_α et H_β par effet Stark, prévu par la théorie (en accord avec l'expérience) est illustré par la figure 83. Les raies H_γ, H_δ éclatent de manière analogue.

6. Le schéma d'éclatement par effet Stark qui vient d'être décrit correspond au cas où *on néglige le spin de l'électron*, donc si on néglige la *structure fine* des raies spectrales. On ne peut le faire que si l'éclatement par effet Stark est notablement plus large que la largeur de la structure fine d'une raie spectrale. Dans des champs électriques de l'ordre de plusieurs dizaines de milliers de V/cm

et au-dessus, la structure fine ne joue pratiquement aucun rôle. On dira qu'il s'agit alors de champs *forts* (comme ce fut le cas pour les champs magnétiques dans l'effet Zeeman). Lorsque l'éclatement par effet Stark est comparable ou

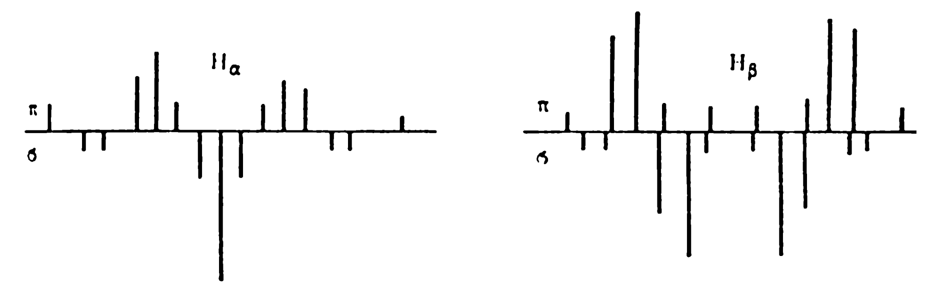


Fig. 83

inférieur à la largeur de la structure fine on dira que le champ électrique appliqué est *faible*. Par conséquent, les résultats donnés ci-dessus se rapportent au cas des champs électriques *forts*. Dans les champs faibles l'effet Stark se complique par la superposition de la *structure fine*.

§ 44. Déplacement de Lamb des niveaux énergétiques des électrons atomiques

1. Il a été indiqué au § 38 que selon la théorie quantique relativiste de Dirac les énergies des atomes et des ions hydrogénoïdes et d'hydrogène ayant les mêmes nombres quantiques n et j devaient être *égales*, quelle que soit la valeur du nombre quantique l. Lorsque j est donné, le nombre l peut prendre les valeurs $j - 1/2$ et $j + +1/2$ suivant le sens du spin de l'électron. Ainsi, au nombre quantique principal $n = 2$ (terme inférieur de la série de Balmer de l'hydrogène) correspondent trois niveaux $2s_{1/2}$, $2p_{1/2}$ et $2p_{3/2}$. Selon la théorie de Dirac les deux premiers niveaux doivent être confondus, ce qui implique qu'ils doivent avoir le même nombre quantique $j=1/2$.

Pour vérifier que les niveaux $2s_{1/2}$ et $2p_{1/2}$ sont effectivement confondus on a cherché à mettre en œuvre les méthodes de la spectroscopie optique en analysant la structure fine de la raie H_α de l'hydrogène, qui résulte des transitions des niveaux avec le nombre quantique principal $n = 3$ aux niveaux avec $n = 2$. Les résultats que l'on obtenait n'étaient pas concordants. Certains chercheurs constataient qu'il y avait un accord parfait entre la structure fine observée et les prévisions de la théorie de Dirac, tandis que d'autres concluaient que les niveaux $2s_{1/2}$ et $2p_{1/2}$ étaient déplacés l'un par rapport à l'autre d'une longueur d'environ 0,03 cm^{-1}, soit de 1000 Hz environ. Cette valeur est prise de 10 fois plus petite que la distance entre ces niveaux et le niveau supérieur $2p_{3/2}$. Les difficultés aux-

quelles se heurtait l'étude de la structure fine étaient dues à ce que les raies spectrales concernées étaient relativement larges et très rapprochées. Les petits écarts observés entre les niveaux $2s_{1/2}$ et $2p_{1/2}$ étaient de l'ordre des erreurs expérimentales. La question fut tranchée en 1947 par Lamb (né en 1913) et Rutherford par la mise en œuvre de *méthodes radiospectroscopiques*. Les déplacements des niveaux de ~1000 Hz auxquels on pouvait s'attendre se situaient dans la gamme des hyperfréquences, ce qui permettait d'utiliser la méthode radiospectroscopique qui permet d'assurer une précision de mesure de 1 Hz environ.

2. L'idée des expériences de Lamb et Retherford se fonde sur ce que le niveau excité $2p_{1/2}$ est instable, tandis que le niveau excité $2s_{1/2}$ est métastable. La durée de vie sur le niveau $2s_{1/2}$ est près de 10^8 fois plus grande que sur le niveau $2p_{1/2}$. En effet la transition radiative du niveau $2p_{1/2}$ au niveau non excité $1s_{1/2}$ avec émission d'un photon est *permise par la règle de sélection* $\Delta l = \pm 1$. Mais la transition $2s_{1/2} \rightarrow 1s_{1/2}$ est *interdite* puisque $\Delta l = 0$. Cette dernière transition ne peut se produire qu'avec émission de *deux photons* et par suite elle se réalise près de 10^8 fois plus lentement que la transition $2p_{1/2} \rightarrow 1s_{1/2}$. En comparaison de la transition $2s_{1/2} \rightarrow 1s_{1/2}$ la transition $2p_{1/2} \rightarrow 1s_{1/2}$ s'effectue 10^8 fois plus rapidement; elle est donc pratiquement instantanée.

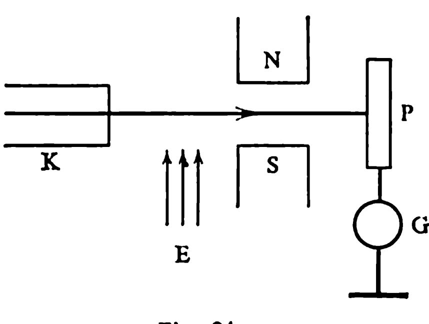


Fig. 84

Dans les expériences de Lamb et Retherford les molécules d'hydrogène étaient thermiquement dissociées dans le four K (fig. 84) et on obtenait un faisceau d'atomes d'hydrogène se trouvant dans l'état fondamental $1s_{1/2}$. Ce faisceau d'atomes était envoyé sur une cible métallique P reliée au galvanomètre G. Comme les atomes du faisceau n'étaient pas excités ils ne pouvaient pas transmettre de l'énergie aux électrons de la cible P et comme les électrons de la cible ne pouvaient être libérés le galvanomètre G n'indiquait aucun courant.

On pouvait cependant exciter une petite partie des atomes (un sur cent millions) et les faire passer dans les états excités $2s_{1/2}$ et $2p_{1/2}$, en réalisant le croisement du faisceau d'atomes avec le faisceau électronique E. L'excitation des atomes étant réalisée par les chocs d'électrons, les règles de sélection pour l'émission *ne sont pas applicables*. L'énergie d'excitation est égale à 10,2 eV. Les atomes se trouvant dans l'état $2p_{1/2}$ ne parviennent pas jusqu'à la cible P,

car ils passent presque instantanément à l'état fondamental $1s_{1/2}$. Seuls les atomes excités à l'état $2s_{1/2}$ peuvent parvenir jusqu'à la cible. En tombant sur la cible, ils transmettent l'énergie d'excitation aux électrons de la cible P, ce qui permet de les libérer. Le galvanomètre se met à enregistrer le passage d'un courant, dont l'intensité permet de se faire une idée du nombre d'atomes se trouvant dans l'état métastable $2s_{1/2}$ qui tombent sur la cible.

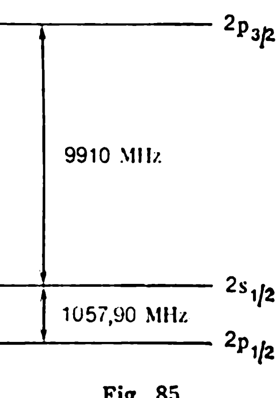


Fig. 85

Faisons agir sur le faisceau d'atomes d'hydrogène, se dirigeant vers la cible, un champ magnétique variant de façon cyclique à une fréquence adéquate ; si les niveaux $2s_{1/2}$ et $2p_{1/2}$ ne sont pas confondus, l'action du champ magnétique provoquera des transitions stimulées entre ces niveaux (cf. § 42. 2). La vitesse de ces transitions stimulées sera maximale à la résonance, lorsque la fréquence du champ magnétique extérieur correspondra à la différence d'énergie entre les niveaux considérés. Passant de l'état $2s_{1/2}$ à l'état $2p_{1/2}$, l'atome retournera presque instantanément à l'état non excité $1s_{1/2}$. Dès lors le nombre d'atomes excités à l'état $2p_{1/2}$ tombant sur la cible commence à diminuer, ce qui se traduit par une diminution du courant passant à travers le galvanomètre G. En notant le minimum du courant, on arrive à déterminer la fréquence de résonance du champ ainsi que la différence d'énergie entre les niveaux $2s_{1/2}$ et $2p_{1/2}$ *).

Les expériences de Lamb et Retherford ont démontré que les niveaux $2s_{1/2}$ et $2p_{1/2}$ sont distincts. La différence en fréquence entre ces niveaux est, pour l'hydrogène, égale à $1057{,}845 \pm 0{,}09$ MHz. Cette différence (et plus généralement toute différence entre les niveaux d'une structure fine ayant les mêmes n et j, mais des l différents) est appelée *déplacement de Lamb* (*Lamb shift*). La figure 85 représente pour l'hydrogène les positions relatives de niveaux $2s_{1/2}$, $2p_{1/2}$, $2p_{3/2}$ avec indication des distances de séparation de ces niveaux. Le déplacement de Lamb est très petit, près de 10 fois plus petit que la distance entre les niveaux $2p_{3/2}$ et $2s_{1/2}$ de la structure fine. Néanmoins, c'est un phénomène très important, dont la théorie

*) On constate que le niveau $2p_{1/2}$ se trouvait au-dessous du niveau $2s_{1/2}$, ce qui implique que les transitions $2s_{1/2} \to 2p_{1/2}$ peuvent se produire en l'absence de tout champ extérieur. Cependant comme l'éclatement est petit et la probabilité d'une transition est proportionnelle au cube de la distance entre les niveaux, la probabilité de la transition envisagée est extrêmement petite (le temps de vie correspondant est de plusieurs années). On peut donc ne tenir aucun compte de cette transition.

a été faite sur la base de l'électrodynamique quantique. Selon cette théorie, le déplacement de Lamb pour l'hydrogène serait égal à 1057,864 ± 0,014 MHz, valeur théorique en excellant accord avec la valeur expérimentale indiquée ci-dessus. Un accord aussi bon a été obtenu pour le deutérium et les atomes d'hélium ionisés une fois.

3. On doit remarquer que le déplacement de Lamb permet de juger dans quelle mesure se trouve vérifiée la loi de Coulomb aux distances atomiques (10^{-8} cm). Supposons que la loi exacte soit de la forme $E \sim 1/r^{2\pm\gamma}$ et que $|\gamma| < 10^{-9}$. Avec ces hypothèses, les calculs auraient donné, pour la variation du déplacement de Lamb, $2s_{1/2}$ — —$2p_{1/2}$, une valeur supérieure à l'erreur de mesure connue. On peut en conclure qu'aux distances atomiques $|\gamma|$ ne saurait être supérieur à 10^{-9} (cf. t. III, § 6, 4). En fait, les écarts à la loi de Coulomb lors de l'interaction d'un électron avec le noyau ne commencent à se manifester que lorsque les forces nucléaires commencent à intervenir. Dans le cas où l'interaction nucléaire est absente chez les particules en interaction (électrons et positons), les études sur les accélérateurs à faisceaux opposés ont permis de prouver que la loi de Coulomb était correcte jusqu'aux distances $\sim 10^{-16}$ cm.

§ 45. Vide physique et interprétation du déplacement de Lamb

1. Bethe (né en 1906) a expliqué le déplacement de Lamb et en a fait les calculs théoriques dans le cadre de l'électrodynamique quantique. Comme celle-ci n'a pas été abordée dans notre cours, on ne peut donner ici qu'un aperçu qualitatif de la théorie du déplacement de Lamb.

Il importe au préalable d'examiner la question du *vide physique* et celle des *particules virtuelles*, ces questions étant les concepts les plus importants de la théorie quantique des champs. Selon la théorie quantique des champs, le vide n'est pas le « néant » absolu où il n'y a rien. Le vide possède de nombreuses propriétés physiques et il peut se trouver dans des états physiques différents. C'est pour cela qu'on lui décerna la dénomination honorifique de « vide physique ». A la différence de l'éther hypothétique du XIX[e] siècle, auquel on attribuait des propriétés mécaniques ne se distinguant pas en principe de celles des milieux matériels ordinaires, la physique moderne s'efforce de définir les propriétés du vide physique uniquement sur la base de faits expérimentaux et de théories physiques dûment confirmés. On n'attribue aucune propriété mécanique au vide physique.

En toute rigueur, on devrait parler de plusieurs vides, selon la nature des particules et des champs qui les concernent. Ainsi, le champ électromagnétique ou champ de photons peut céder son énergie par quanta $h\nu$. Chaque fois qu'un quantum est cédé, le

nombre de photons diminue d'une unité. A la suite d'une succession de tels processus, on arrive finalement à un état où le nombre de quanta contenus dans le système devient égal à zéro. Mais à la différence des conceptions classiques, le champ électromagnétique ne disparaît pas pour autant et passe à un état de plus petite énergie qu'il est impossible d'enlever au champ. Cette assertion résulte de l'existence d'une énergie au point zéro, donc au bout du compte, du principe d'incertitude. L'état du champ électromagnétique, possédant la plus petite énergie possible et ne comportant aucun photon, est appelé *état de vide du champ électromagnétique* ou *vide photonique*. Le champ électromagnétique à l'état de vide ne peut céder aucune énergie, mais cela ne signifie pas qu'il ne peut se manifester d'aucune façon. Ce champ peut être à l'origine des différents phénomènes physiques observables dont il sera question plus loin.

Pour les autres particules aussi on introduit la notion du vide en tant qu'un état d'énergie minimale du champ des particules correspondantes. On définit ainsi le *vide des électrons-positons* dont les particules électrons et positons ne se distinguent les uns des autres que par les signes des charges électriques. Il existe aussi le *vide des mésons* π, etc. Dans l'étude des interactions des champs, on peut désigner par le terme « vide » l'état énergétique le plus bas du système de ces champs.

Si on fournit une énergie suffisante à un champ se trouvant dans l'état de vide, celui-ci sera excité, ce qui signifie qu'il donnera naissance à une particule — quantum de ce champ. Par exemple, le vide peut être le siège de la création de paires électron-positon. On peut décrire la création d'une particule comme la transition de l'état de vide « inobservable » à un état réel.

2. Selon les conceptions modernes, l'interaction des particules s'effectue grâce à un échange d'autres particules auxquelles correspond leur propre vide. Par exemple, les interactions électromagnétiques entre des particules portant des charges électriques s'effectuent par l'intermédiaire du vide électromagnétique. L'une des charges électriques émet un photon qu'absorbe l'autre charge et celle-ci émet à son tour un photon qu'absorbe la première charge. Ainsi les charges électriques échangent des photons, en conséquence de quoi l'état de zéro (i.e. l'état non excité) du vide se trouve modifié ; selon les conceptions modernes, cela se manifeste dans la force d'interaction des particules échangeant des photons. De même, les interactions entre les nucléons (protons et neutrons) s'effectuent par l'intermédiaire du vide des mésons π. L'un des nucléons émet un méson π, l'autre l'absorbe et vice versa. Telle est la conception moderne des forces nucléaires.

A première vue, ces conceptions de l'interaction paraissent être en contradiction avec ce qui avait été dit au § 1, 7. Pour fixer les idées, considérons l'interaction de charges électriques, mais les

conclusions qu'on en tirera seront valables pour tous les autres types d'interactions. Avant que commence l'interaction, chacune des particules est libre et aucune particule libre ne peut ni émettre ni capter un quantum. S'il n'en était pas ainsi, dans les processus d'émission et d'absorption d'un quantum d'énergie, la loi de la conservation de l'énergie ou la loi de la conservation de l'impulsion serait en défaut. Mais tout cela ne concerne que l'émission de vraies particules. Or l'interaction est assurée par des *particules virtuelles* et de ce fait la contradiction qui vient d'être évoquée se trouve levée. On aurait peut-être dû parler non pas de particules échangées, mais de particules virtuelles échangées lors des interactions.

Les particules virtuelles n'existent que dans des états intermédiaires et pendant des intervalles de temps très courts, ce qui empêche leur enregistrement expérimental. La durée d'existence Δt d'une particule virtuelle est liée à l'incertitude sur son énergie par la relation $\Delta t \Delta \mathcal{E} \geqslant \hbar$; on peut écrire aussi $\Delta x \Delta p \geqslant \hbar$. Compte tenu de ces relations d'incertitude, les lois de la conservation de l'énergie ne s'opposent plus à l'émission de quanta par des vraies particules libres, à condition que ces quanta possèdent l'énergie $\Delta \mathcal{E}$ et qu'ils existent pendant un intervalle de temps $\Delta t \sim \hbar / \Delta \mathcal{E}$. On peut dire qu'une particule virtuelle est une particule pour laquelle n'est pas vérifiée la relation classique ordinaire entre l'impulsion et l'énergie

$$\mathcal{E}^2 = (pc)^2 + m_0^2 c^4.$$

3. Pour le temps Δt, il est tout naturel de prendre la durée de l'acte d'échange entre les particules se trouvant en interaction. Si on admet encore que l'interaction se propage avec la vitesse maximale C, on peut écrire $\Delta t = L/C$, L étant le rayon de l'interaction transportée. L'énergie du quantum est donc égale à $\Delta \mathcal{E} = m_0 c^2 + \hbar c / L$, m_0 étant la masse de repos du quantum. Il s'ensuit que

$$L = \hbar / (m_0 c)^2. \qquad (45.1)$$

Plus la masse m_0 du quantum transportant l'interaction est grande, plus le rayon d'action des forces en cause est petit. Cette proposition sera développée dans le volume consacré à la physique nucléaire, et pour l'instant on indiquera seulement que pour le photon $m_0 = 0$. Il s'ensuit que le rayon d'action des forces coulombiennes est infini.

En conséquence, le principe d'incertitude impose qu'on doit admettre que chaque particule chargée est entourée d'un nuage de photons virtuels émis et absorbés. Les autres particules sont également entourées de quanta correspondants assurant le transport de l'interaction. Cette conception satisfait aux vues de Faraday et de Maxwell qui jugeaient inacceptable l'idée de l'action directe à distance. L'interaction doit faire intervenir un agent intermédiaire. Mais on ne peut concrétiser la nature de cet agent comme on le faisait

du temps de l'éther universel classique, qui n'est plus qu'une réminescence historique.

4. Le vide est une superposition des vibrations de zéro du champ, i.e. d'états donnant lieu à la création et à l'annihilation des photons virtuels, de paires d'électrons-positons virtuels, ainsi que d'autres paires de particules et d'antiparticules virtuelles. Ces particules virtuelles sont en interactions mutuelles et avec des particules vraies. Un photon virtuel nouvellement créé peut donner naissance à une paire électron-positon. L'annihilation de cette dernière peut s'accompagner de la création de nouveaux photons virtuels, etc. La distribution des électrons et des positons virtuels dans un champ électrique extérieur, par exemple dans le champ du noyau atomique, n'est pas uniforme. Les positons virtuels se déplacent de préférence dans le sens du champ électrique et les électrons virtuels en sens contraire. Le résultat en est un phénomène appelé *polarisation du vide*. Ce phénomène est analogue à la polarisation ordinaire des diélectriques placés dans un champ extérieur, à la différence près que dans ce dernier cas il s'agit du déplacement de charges électriques réelles, tandis que dans le vide il s'agit du déplacement de charges virtuelles.

5. Passons maintenant à l'explication du déplacement de Lamb. La principale contribution ($\sim\alpha^3 R$, α étant la constante de la structure fine et R la constante de Rydberg) est due à deux effets de vide (appelés aussi corrections radiatives). Premièrement, c'est l'émission et l'absorption de photons virtuels par l'électron lié, ce qui affecte sa masse effective et conduit à l'apparition d'un moment magnétique anomal. Deuxièmement, c'est une polarisation du vide, correspondant à la création et à l'annihilation dans le vide de paires électron-positon virtuelles, ce qui se traduit par une déformation du potentiel coulombien du noyau à une distance de l'ordre de la longueur d'onde de Compton de l'électron ($\sim 10^{-11}$ cm). Comme la longueur d'onde de Compton de l'électron est beaucoup plus petite que le rayon moyen des orbites de Bohr de l'atome d'hydrogène, le déplacement de Lamb dans l'hydrogène est surtout dû à la première cause (variation de la masse effective de l'électron). La polarisation du vide provoque un déplacement égal de tous les niveaux. La contribution de la polarisation du vide au déplacement de Lamb des niveaux est petite (dans l'hydrogène elle représente 3 % du déplacement total du niveau fondamental).

Conformément à la théorie qui est corroborée par l'expérience, le déplacement de Lamb du niveau d'énergie est proportionnel à la quatrième puissance du numéro atomique et inversement proportionnel à la troisième puissance du nombre quantique principal.

Un raisonnement simple permet de faire comprendre l'influence qu'exerce le nombre quantique principal sur le déplacement de Lamb. On peut dire, pour faire image, que l'électron *s* séjourne la plupart du temps auprès du noyau, où le champ électrique est in-

tense et fortement inhomogène, tandis que l'électron p se trouve en moyenne pendant plus de temps à des distances plus grandes, là où le champ est moins fort et moins inhomogène. D'autre part, l'interaction avec le vide (émission et absorption de photons virtuels) fait trembler, si on peut dire, l'électron. En termes de la physique classique, la trajectoire de l'électron n'est pas lisse (un cercle par exemple) mais sinueuse. L'électron tantôt s'éloigne du noyau, tantôt s'en rapproche, mais de façon désordonnée. L'énergie potentielle de l'électron dans le champ du noyau $U \sim 1/r$. Lorsque r augmente de Δr, l'énergie U varie de

$$\Delta U \sim \frac{1}{r+\Delta r} - \frac{1}{2} = -\frac{\Delta r}{r(r+\Delta r)}$$

et lorsque r diminue de Δr

$$\Delta U' \sim \frac{\Delta r}{r(r-\Delta r)},$$

$\Delta U'$ étant plus grand en valeur absolue que ΔU. Cela signifie que le tremblement de l'électron dans le vide donne lieu à un changement du signe de la variation de son énergie potentielle U. A proximité du noyau, la variation de l'énergie potentielle est particulièrement grande, puisque U y est grand et varie fortement avec la distance. Ainsi, les contributions à l'énergie totale $\mathcal{E}$ des électrons s et p sont inégales, celle de l'électron s étant prépondérante. C'est ce qui provoque une différence entre les énergies des niveaux des électrons s et p, qui coïncideraient sans cela (compte non tenu du déplacement des niveaux dû à la polarisation du vide).

CHAPITRE VI

SYSTÈMES ATOMIQUES À PLUSIEURS ÉLECTRONS

§ 46. Principe d'indiscernabilité des particules identiques. Principe de Pauli

1. A la différence des corps macroscopiques, les particules de même espèce du microcosme (électrons, protons, neutrons, toutes les particules dites élémentaires, atomes, etc.) possèdent des propriétés *absolument identiques*: elles ont même masse, même charge électrique, même spin, etc. On doit se demander comment on pourrait distinguer deux particules de même espèce. Considérons, par exemple, un système de deux électrons. Repérons à un certain instant (pris pour instant initial) les positions des électrons en les notant *1* et *2*. Selon des conceptions classiques, tout électron se meut suivant une trajectoire déterminée, ce qui permettrait de suivre les mouvements de chacun de ces électrons. Ayant décelé l'électron à un instant ultérieur, on pourrait, en principe, décider de quel électron il s'agit, le numéro *1* ou le numéro *2*. En intervertissant les positions et les vitesses des deux électrons, on obtient un *nouvel état* du système. Quoique ce nouveau système possède exactement les mêmes propriétés que le système initial, il s'en distingue par la numérotation des électrons. Au point de vue classique, les particules identiques seraient, *en principe*, *discernables* et *conserveraient leurs identités*.

La situation est radicalement différente en mécanique quantique qui réfute la conception classique du mouvement des particules sur des trajectoires. En mécanique quantique, l'état d'un système de particules est décrit par une fonction d'onde que l'on interprète de façon probabiliste. Dans notre exemple, la fonction d'onde dépend du temps et des coordonnées des deux électrons. Si on décèle à un certain instant l'un des électrons, *on ne peut, en principe, décider* s'il s'agit de l'électron *1* ou de l'électron *2*. L'impossibilité de trouver une solution de principe à une certaine question signifie que la question est *mal* posée. Au point de vue de la mécanique-quantique, la question de discernabilité de deux particules identiques est un faux problème. Si on permute deux particules identiques, l'*expérience ne permet pas de mettre en évidence* le résultat de cette permutation. Il est naturel d'exiger de toute théorie qu'elle considère deux états que l'on ne peut discerner par aucune expérience, comme *un seul et même état*. C'est ce que fait la mécanique quantique qui accepte que la permutation de deux particules identiques ne crée pas un nouvel

état : l'état du système reste exactement le même après permutation. Les particules identiques sont *indiscernables par principe* et sont donc *dénuées d'identité*. On ne peut parler de l'état d'un système de particules identiques qu'en tant que *tout* et non pas de l'état de chaque particule prise individuellement. Cette hypothèse peut être formulée sous la forme du *principe d'indiscernabilité de particules identiques*.

Dans un système de particules identiques ne se réalisent que les états qui ne changent pas si on permute deux particules quelconques.

Ce principe de la mécanique quantique est essentiellement nouveau en ce sens qu'il ne découle pas de ses autres propositions fondamentales ; on démontre que ce nouveau principe ne contredit nullement les autres propositions de la théorie. On doit accepter ce principe pour la raison qu'il est confirmé par toutes les données factuelles.

2. En mécanique quantique, on caractérise l'état d'un système de particules par la fonction d'onde. Il s'agit de savoir quelles fonctions d'onde sont acceptables en ce qu'elles satisfont au principe d'indiscernabilité de particules identiques. Pour trouver une réponse à cette question, il suffit d'envisager un système de deux particules identiques. En faisant abstraction de la dépendance de la fonction d'onde avec le temps t, on peut l'écrire sous la forme $\psi(q_1, q_2)$. S'il s'agit de particules sans spin, on entendra par q_1 l'ensemble des trois coordonnées spatiales de la première particule et par q_2 l'ensemble des coordonnées spatiales de la seconde particule. Dans le cas de particules à spin, on doit ajouter à la triade des coordonnées spatiales les *coordonnées de spin* qui peuvent prendre une suite de valeurs discrète. Par exemple, pour les électrons, la projection du spin peut prendre deux valeurs $+1/2$ et $-1/2$. Ce sont ces valeurs qui constituent, dans ce cas particulier, les coordonnées de spin.

Si on permute les particules *1* et *2*, la fonction d'onde sera $\psi(q_2, q_1)$. Cette opération peut être considérée comme l'action d'un opérateur linéaire $\hat{P}$, dit *opérateur de permutation*, sur la fonction $\psi(q_1, q_2)$:

$$\psi(q_2, q_1) = \hat{P}\psi(q_1, q_2).$$

En permutant les particules encore une fois, on retrouve la fonction initiale $\psi(q_1, q_2) = \hat{P}\psi(q_2, q_1) = \hat{P}^2\psi(q_1, q_2)$.

On en déduit que $\hat{P}^2 = 1$ et par suite $\hat{P} = \pm 1$. On peut donc utiliser deux types de fonctions d'onde :

$$\psi_s(q_1, q_2) = \psi_s(q_2, q_1) \quad (46.1)$$

et

$$\psi_a(q_1, q_2) = -\psi_a(q_2, q_1). \quad (46.2)$$

Dans le premier cas, la fonction d'onde ne change pas à la suite d'une permutation des particules, et on l'appelle fonction d'onde *symétrique* (l'indice s le rappelle). Dans le second cas, la fonction

est appelée *antisymétrique* et on l'affecte de l'indice a. Lorsqu'on permute des particules identiques, la fonction d'onde antisymétrique change de signe.

Ces résultats sont valables pour les systèmes composés d'un nombre quelconque de particules identiques. Pour ces systèmes la symétrie ou l'antisymétrie de la fonction d'onde doit se manifester dès qu'on permute deux particules identiques quelconques du système. Il est formellement admissible que le système comportant plus de deux particules a les propriétés de symétrie plus compliquées. La limitation complémentaire par les fonctions symétriques et antisymétriques est dictée par l'expérience.

3. Les particules dont l'état est décrit par des fonctions d'onde symétriques sont appelées *particules de Bose* ou *bosons*, et celles dont l'état est décrit par des fonctions d'onde antisymétriques sont appelées *particules de Fermi* ou *fermions*. Ces dénominations résultent de ce que les systèmes composés de bosons sont régis par la *statistique de Bose-Einstein* et ceux composés de fermions obéissent à la *statistique de Fermi-Dirac* (cf. t. II, § 82). Au nombre des bosons on trouve les photons, les mésons π et K, en général toutes les particules à spin nul ou entier. Les électrons, les protons, les neutrons, les neutrinos et toutes les autres particules et antiparticules élémentaires à spin demi-entier sont des fermions. Le lien existant entre le spin et la statistique fut établi d'abord empiriquement pour les photons et les électrons, puis en 1940 fut étendu sur une base théorique par Pauli à toutes les particules et antiparticules élémentaires. Pauli établit cette relation en partant des principes les plus généraux de la théorie quantique : l'invariance relativiste, la non-négativité de l'énergie totale, la causalité, etc.

Cette relation entre le spin et la statistique reste valable pour les particules complexes construites avec des particules élémentaires, par exemple pour les noyaux atomiques, les atomes et les molécules, à condition que les phénomènes concernés ne font intervenir que des énergies suffisamment faibles pour que chacune de ces particules complexes se comporte comme un tout. L'appartenance d'une particule complexe au groupe des bosons ou à celui des fermions dépend de la valeur de son spin.

Par exemple, l'atome d'hydrogène se compose de deux particules de Fermi : le proton et l'électron, chacun ayant un spin égal à 1/2. Le moment mécanique total, i.e. le spin de l'atome d'hydrogène dans l'état normal, peut être égal soit à 0 (les spins du proton et de l'électron sont antiparallèles), soit à 1 (spins parallèles). Dans les deux cas l'atome d'hydrogène à l'état normal est un boson.

Considérons un autre exemple. Le noyau de l'atome d'hélium ${}^4_2He$, i.e. la particule α, se compose de deux protons et de deux neutrons. Le spin de ce noyau est égal à 0, ce qui signifie que la particule α est un boson. L'atome d'hélium ${}^4_2He$ à l'état normal est égale-

ment un boson. Mais le noyau $^3_2$He se compose d'un nombre impair (3) de particules ayant des spins 1/2 (deux protons et un neutron). Comme le spin de ce noyau est impair, c'est un fermion. L'atome $^3_2$He lui-même est un fermion. Les noyaux et les atomes $^4_2$He se conforment à la statistique de Bose-Einstein, tandis que les noyaux et les atomes $^3_2$He se conforment à la statistique de Fermi-Dirac. Cette différence se manifeste en ce qu'à une température proche du zéro absolu $^4_2$He est superfluide, tandis que $^3_2$He ne l'est pas (pour plus de détails voir § 61, 7).

4. Appliquons les principes de symétrie et d'antisymétrie à un système de particules identiques non interagissantes. Commençons par raisonner comme si les particules étaient dénuées de spin (dans ce cours on n'étudiera pas les équations correspondant aux fonctions d'onde de particules à spin), pour étendre ensuite les résultats obtenus au cas de particules à spin.

Envisageons le cas le plus simple d'un système composé de deux particules identiques dont les coordonnées cartésiennes seront notées x_1, y_1, z_1 et x_2, y_2, z_2. En notant $\hat{H}$ l'opérateur de Hamilton du système tout entier et $\hat{H}_1$, $\hat{H}_2$ les opérateurs de chacune des particules, on a $\hat{H} = \hat{H}_1 + \hat{H}_2$. Pour un système de particules se trouvant dans des états stationnaires, l'équation de Schrödinger s'écrit

$$\hat{H}\psi \equiv (\hat{H}_1 + \hat{H}_2)\,\psi = \mathscr{E}\psi. \tag{46.3}$$

Comme les particules n'interagissent pas entre elles, l'opérateur $\hat{H}_1$ ne dépend que des coordonnées de la première particule et l'opérateur $\hat{H}_2$ que des coordonnées de la deuxième particule, à savoir

$$\begin{aligned}\hat{H}_1 &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}\right) + U\,(x_1,\ y_1,\ z_1),\\ \hat{H}_2 &= -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}\right) + U\,(x_2,\ y_2,\ z_2).\end{aligned} \tag{46.4}$$

Les deux opérateurs $\hat{H}_1$ et $\hat{H}_2$ sont exactement les mêmes puisque les deux particules sont identiques, mais ils dépendent de coordonnées différentes.

Pour résoudre l'équation (46.3) on utilisera la méthode qui a été déjà utilisée au § 26. Puisque l'opérateur $\hat{H}_1$ n'agit que sur les coordonnées x_1, y_1, z_1, et l'opérateur $\hat{H}_2$ que sur les coordonnées x_2, y_2, z_2, l'équation (46.3) se scinde en deux équations :

$$\hat{H}_1\psi = \mathscr{E}_1\psi, \quad \hat{H}_2\psi = \mathscr{E}_2\psi, \tag{46.5}$$

où $\mathscr{E}_1$ et $\mathscr{E}_2$ sont des constantes vérifiant l'égalité $\mathscr{E}_1 + \mathscr{E}_2 = \mathscr{E}$. La solution de la première équation (46.5) est de la forme $\varphi_\beta\varphi_\alpha\,(1)$,

où (*1*) dénote l'ensemble des coordonnées x_1, y_1, z_1 de la première particule; on dénote par (*2*) l'ensemble des coordonnées x_2, y_2, z_2 de la seconde particule. Dans le cas le plus général, le facteur φ_β peut dépendre des coordonnées *2*, i.e. $\varphi_\beta = \varphi_\beta(2)$. La fonction $\varphi_\beta(2)$ doit être choisie de façon que la deuxième équation (46.5) soit vérifiée. Il faut donc que

$$\hat{H}_1\varphi_\alpha(1) = \mathscr{E}_1\varphi_\alpha(1) \quad \text{et} \quad \hat{H}_2\varphi_\beta(2) = \mathscr{E}_2\varphi_\beta(2). \qquad (46.6)$$

La fonction $\varphi_\alpha(1)$ décrit donc l'état de la première particule d'énergie $\mathscr{E}_1$, et la fonction $\varphi_\beta(2)$ l'état de la seconde particule d'énergie $\mathscr{E}_2$. La relation $\mathscr{E} = \mathscr{E}_1 + \mathscr{E}_2$ signifie que l'énergie d'un système de deux particules n'interagissant pas entre elles est égale à la somme des énergies de ces particules, comme il fallait s'y attendre. La solution de l'équation (46.3) est donc de la forme $\psi = \varphi_\alpha(1)\varphi_\beta(2)$, qui est une solution à *variables séparables.*

Les équations (46.6) sont en fait identiques, puisqu'elles ne diffèrent l'une de l'autre que par les notations des variables indépendantes et par les valeurs des constantes $\mathscr{E}_1$ et $\mathscr{E}_2$. Bien entendu, à la valeur propre $\mathscr{E}_1$ peuvent correspondre plusieurs solutions indépendantes de la première équation (46.6). Dans le cas général, on doit entendre par $\varphi_\alpha(1)$ une superposition linéaire de ces solutions avec des coefficients constants, cette superposition étant aussi solution de l'équation (46.3). Il en est de même de la fonction $\varphi_\beta(2)$.

Avant d'aller plus loin, cherchons à tenir compte du spin. Pour cela il suffit d'entendre par (*1*) et (*2*) l'ensemble des coordonnées spatiales et des *coordonnées de spin* des particules. On déterminera de la même façon que ci-dessus la fonction $\varphi_\alpha(1)\varphi_\beta(2)$ qui est solution de l'équation se substituant à l'équation de Schrödinger dans le cas où les particules possèdent des spins (cette équation est connue sous le nom d'équation de Pauli). Comme les particules sont identiques, la fonction $\varphi_\alpha(2)\varphi_\beta(1)$ est solution de cette même équation. Cette fonction se déduit de la fonction $\varphi_\alpha(1)\varphi_\beta(2)$ si on permute les particules *1* et *2*, ce qui revient à permuter leurs coordonnées spatiales et de spin.

Or aucune de ces fonctions ne saurait satisfaire au principe de symétrie ou d'antisymétrie, mais on peut en former des combinaisons linéaires à coefficients constants qui seront aussi des solutions d'une équation du type de l'équation de Schrödinger. Parmi les différentes combinaisons, on trouve une *fonction symétrique*

$$\psi_s = \varphi_\alpha(1)\varphi_\beta(2) + \varphi_\alpha(2)\varphi_\beta(1) \qquad (46.7)$$

et une *fonction antisymétrique*

$$\psi_a = \varphi_\alpha(1)\varphi_\beta(2) - \varphi_\alpha(2)\varphi_\beta(1). \qquad (46.8)$$

L'état décrit par la fonction ψ_s peut se réaliser effectivement dans la

Nature pour un système de deux bosons identiques, et l'état décrit par la fonction ψ_a dans le cas de deux fermions identiques.

5. Ces raisonnements se laissent étendre sans aucune difficulté aux systèmes composés d'un nombre quelconque N de particules identiques. On obtient d'abord une fonction d'onde telle que

$$\psi = \psi_\alpha(1)\,\psi_\beta(2)\,\psi_\gamma(3)\ldots\psi_\omega(N)$$

à variables séparables où on a gardé les mêmes notations que ci-dessus. En permutant successivement les particules deux par deux, on obtient de nouvelles fonctions d'onde à variables séparables:

$$\psi' = \psi_\alpha(2)\psi_\beta(1)\,\psi_\gamma(3)\ldots\psi_\omega(N),$$
$$\psi'' = \psi_\alpha(3)\psi_\beta(2)\,\psi_\gamma(1)\ldots\psi_\omega(N)$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

auxquelles correspond la même valeur de $\mathscr{E}$. L'état caractérisé par cette valeur de $\mathscr{E}$ est N fois dégénéré. Néanmoins aucun de ces états n'est réalisé dans la Nature.

Dans le cas de bosons, ne se réalise que la combinaison symétrique

$$\psi_s = \sum \hat{P}\psi, \tag{46.9}$$

où $\hat{P}$ est l'opérateur de permutation de deux particules, la somme étant étendue à toutes les permutations possibles, y compris la permutation identique ne modifiant les coordonnées d'aucune particule et qui ne modifie pas la fonction ψ. Ainsi, la somme (46.9) comporte N termes.

Dans le cas de fermions, c'est la combinaison antisymétrique

$$\psi_a = \sum \pm \hat{P}\psi \tag{46.10}$$

qui se trouve réalisée; au signe plus correspond un nombre pair de transpositions (permutations de deux particules) et au signe moins un nombre impair de transpositions. Il s'ensuit que la fonction antisymétrique peut être présentée sous la forme d'un déterminant

$$\psi_a = \begin{vmatrix} \psi_\alpha(1) & \psi_\beta(1) & \ldots & \psi_\omega(1) \\ \psi_\alpha(2) & \psi_\beta(2) & \ldots & \psi_\omega(2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \psi_\alpha(N) & \psi_\beta(N) & \ldots & \psi_\omega(N) \end{vmatrix}. \tag{46.11}$$

En règle générale, on normalise les fonctions d'onde (46.10) et (46.11) en les multipliant par des coefficients numériques déterminés. Dans le cas qui nous occupe ici, la normalisation ne jouant aucun rôle, on ne la fera pas.

6. Lorsqu'il s'agit d'un système de particules identiques qui n'interagissent pas entre elles, il convient de parler aussi bien de

l'état du système tout entier que de l'*état d'une seule particule.* On peut dire, par exemple, qu'une certaine particule se trouve dans l'état ψ_α et une autre dans l'état ψ_β. Les fonctions d'onde des fermions se comportent autrement que celles des bosons.

Dans un système de fermions identiques il ne peut y avoir plus de deux particules se trouvant dans le même état. Dans le cas particulier d'un système ne comportant que deux particules, cette assertion découle directement de la formule (46.8) puisque si $\psi_\alpha \equiv \psi_\beta$, la fonction d'onde ψ_a s'annule, ce qui physiquement ne peut correspondre à aucun état. Dans le cas général de N particules, ce résultat ressort de la formule (46.10) et plus particulièrement de la formule (46.11), puisque pour $\psi_\alpha \equiv \psi_\beta$ on obtiendrait un déterminant avec deux colonnes identiques, ce qui signifie que le déterminant est égal à zéro.

Par conséquent, *dans un système de fermions identiques on ne peut trouver deux particules se trouvant dans le même état.* Cette proposition exprime le *principe d'interdiction de Pauli* qui l'énonça pour les électrons appartenant aux atomes avant que fut élaborée la mécanique quantique. Dans la formulation initiale, ce principe affirmait qu'*aucun atome ne peut comporter deux électrons dont les quatre nombres quantiques soient les mêmes.* Le principe de Pauli est moins précis et moins clair que le *principe d'antisymétrie des fonctions d'onde.* Ce dernier reste valable dans le cas de particules en interactions mutuelles, tandis que le principe de Pauli ne concerne que les états de particules isolées dont on ne peut parler en toute rigueur qu'en l'absence de toute interaction. Néanmoins le principe de Pauli, même dans sa formulation initiale, rendit beaucoup de services, notamment pour la justification du système périodique de Mendéléev (1834-1907), ainsi que de certaines régularités de la structure des spectres. On appelle parfois le principe d'antisymétrie des fonctions d'onde des fermions *principe de Pauli généralisé.*

En ce qui concerne les bosons, le *principe de symétrie des fonctions d'onde* n'impose à leurs états aucune limitation semblable au principe de Pauli. Un nombre *quelconque* de bosons identiques peuvent se trouver dans un seul et même état. Ce résultat découle directement des expressions (46.7) et (46.9) des fonctions d'onde des bosons.

§ 47. Interprétation physique du système périodique des éléments chimiques de Mendéléev

1. La répétition périodique des propriétés des éléments chimiques, qui a été découverte par D. I. Mendéléev en 1869 et qui trouve son expression dans le *système périodique*, a été interprétée en 1922 par Bohr sur la base de sa théorie de la structure de l'atome. Il s'avéra que ce n'est pas la masse atomique (poids atomique selon l'ancienne terminologie), mais la *charge du noyau* qui détermine la pé-

riodicité des propriétés des éléments chimiques. En prenant pour unité de charge la charge élémentaire *e*, la charge du noyau s'exprime par un nombre entier que l'on dénote par *Z*. C'est le nombre *Z* qui détermine le numéro de l'élément chimique dans le tableau périodique. C'est pour cette raison qu'on l'appelle *numéro d'ordre* de l'élément. La charge du noyau est numériquement égale au nombre d'électrons gravitant autour du noyau de l'atome neutre. Les propriétés des éléments dépendent en premier lieu du nombre d'électrons contenus dans l'enveloppe électronique de l'atome, ainsi que de la structure de cette enveloppe. En ce qui concerne les *propriétés chimiques* des éléments, elles sont déterminées par les *électrons périphériques* de l'enveloppe électronique. C'est après que Pauli formula son principe, en 1925, que l'interprétation du système périodique fit un grand pas en avant. L'application de ce principe permit de déterminer le nombre maximal d'électrons pouvant se trouver dans des états quantiques déterminés dans l'enveloppe électronique de l'atome. On comprit aussitôt la répartition des éléments chimiques par groupes et par périodes du système périodique que Mendéléev dressa empiriquement et que d'autres chimistes complétèrent par la suite.

Le système périodique des éléments est représenté par le tableau 5. On trouve à l'état naturel 90 éléments chimiques. L'uranium est l'élément ayant le plus grand numéro atomique (92). Les éléments $_{43}Tc$ (technétium) et $_{61}Pm$ (prométhium) n'existent pas à l'état naturel. Ce sont des éléments radioactifs dont la période de désintégration est beaucoup plus petite que l'âge de notre planète (la période de l'isotope ^{99}Tc, qui est le plus important, est de $2{,}12 \cdot 10^5$ années, et les périodes de Pm sont comprises entre quelques années et plusieurs dizaines d'années). Depuis que la Terre s'est formée, ces éléments se sont complètement désintégrés. Ils ne pouvaient pas se former par transmutation des noyaux ayant des charges proches (parmi ceux qui se rencontrent dans la Nature), car tous ces noyaux sont stables. Les éléments dont le numéro atomique est supérieur à 92 sont appelés éléments *transuraniens*, tous ces éléments sont radioactifs et ont été produits artificiellement. On a réussi à produire l'élément transuranien ayant le numéro 109.

Les procédés d'obtention des éléments transuraniens seront brièvement décrits dans le volume consacré à la physique nucléaire. Pour l'instant, signalons qu'en 1983 on a fait savoir que les savants de Darmstadt (R.F.A.) avaient obtenu 1 atome de l'élément n° 109. En 1984, à Doubna, on a réussi à obtenir 8 atomes de cet élément. La même année, toujours à Doubna, on a réussi à faire la synthèse de deux isotopes de l'élément 108. Les dénominations nobélium (pour l'élément 102) et lawrencium (pour l'élément 103) ne sont pas universellement reconnues pour la raison que la communication initiale annonçant leur synthèse et leurs noms de baptême

Tableau 5

	I	II	III	IV	V	VI	VII	VIII			0
1	1 H 1,0080										2 He 4,0026
2	3 Li 6,939	4 Be 9,0122	5 B 10,811	6 C 12,011	7 N 14,0067	8 O 15,9994	9 F 18,9984				10 Ne 20,183
3	11 Na 22,9898	12 Mg 24,312	13 Al 26,9815	14 Si 28,086	15 P 30,9738	16 S 32,064	17 Cl 35,453				18 Ar 39,948
4	19 K 39,102	20 Ca 40,08	21 Sc 44,956	22 Ti 47,90	23 V 50,942	24 Cr 51,996	25 Mn 54,938	26 Fe 55,847	27 Co 58,9332	28 Ni 58,71	
	29 Cu 63,54	30 Zn 65,37	31 Ga 69,72	32 Ge 72,59	33 As 74,9216	34 Se 78,96	35 Br 79,909				36 Kr 83,80
5	37 Rb 85,47	38 Sr 87,62	39 Y 88,905	40 Zr 91,22	41 Nb 92,906	42 Mo 95,94	43 Tc 99	44 Ru 101,07	45 Rh 102,9	46 Pd 106,4	
	47 Ag 107,87	48 Cd 112,40	49 In 114,82	50 Sn 118,69	51 Sb 121,75	52 Te 127,60	53 I 126,9044				54 Xe 131,30
6	55 Cs 132,905	56 Ba 137,34	57 La* 138,92	72 Hf 178,49	73 Ta 180,948	74 W 183,85	75 Re 186,2	76 Os 190,2	77 Ir 192,2	78 Pt 195,09	
	79 Au 196,967	80 Hg 200,59	81 Tl 204,37	82 Pb 207,19	83 Bi 208,98	84 Po 210	85 At 210				86 Rn 222
7	87 Fr 223	88 Ra 226,05	89 Ac** 227	104 Ku (261)	105 Ns	106	107	109			

*58—71 Série des lanthanides	58 Ce 140,12	59 Pr 140,97	60 Nd 144,24	61 Pm 150,35	62 Sm 150,35	63 Eu 151,96	64 Gd 157,25	65 Tb 158,924	66 Dy 162,50	67 Ho 164,93	68 Er 167,26	69 Tm 168,934	70 Yb 173,04	71 Lu 174,97
**90—98 Série des actinides	90 Th 232,038	91 Pa 231	92 U 238,03	93 Np 237	94 Pu (242)	95 Am (243)	96 Cm (247)	97 Bk (249)	98 Cf (251)	99 Es (254)	100 Fm (257)	101 Md (258)	102 Jl (255)	103 Rf (256)

n'a pas été confirmée ultérieurement. Les savants de Doubna ont suggéré de dénommer joliotium (Jl) l'élément 102, en l'honneur de Joliot-Curie et rutherfordium (Rf) l'élément 103, en l'honneur de Rutherford. Ils ont suggéré aussi d'appeler nielsbohrium (Ns) le 105[e] élément, en l'honneur de Niels Bohr. Les autres éléments transuraniens n'ont pas encore reçu de dénominations. En prenant exemple sur D. I. Mendéléev, on pourrait les dénommer: ekatungstène, ekarhénium, ekaosmium, ekairidium, le préfixe eka signifiant que les éléments correspondants se placent respectivement au-dessous du tungstène, du rhénium, de l'osmium et de l'iridium.

2. La répétition périodique des propriétés des éléments chimiques est la manifestation visible de la structure interne des enveloppes électroniques des atomes. L'élaboration de la théorie du système périodique des éléments se ramène à l'étude de la structure de l'enveloppe électronique.

Dans la théorie moderne du système périodique, on cherche à caractériser les états des différents électrons *pris individuellement* plutôt que l'état de l'enveloppe électronique toute entière. C'est assurément un défaut majeur de la théorie. L'état de l'enveloppe électronique d'un atome devrait être caractérisé par la fonction d'onde de toutes les coordonnées des électrons qui en font partie; dans le cas d'interactions mutuelles des électrons, ces coordonnées ne peuvent être séparées, ce qui signifie que cette fonction d'onde ne pourrait pas être représentée sous forme d'un produit des fonctions d'onde des électrons individuels. Il s'ensuit que c'est seulement l'état de l'enveloppe électronique toute entière qui peut avoir une signification exacte, la connaissance des états des électrons individuels étant insuffisante. Mais le problème de la détermination exacte de la fonction d'onde d'un système de plusieurs particules en interactions mutuelles est encore plus inaccessible aux mathématiciens que le fameux problème de plusieurs corps (notamment des trois corps) de la mécanique classique. C'est pour cela qu'on est réduit à utiliser une représentation approchée des états des électrons individuels.

L'état d'un électron individuel est caractérisé par un ensemble de quatre nombres quantiques, le plus commode étant l'ensemble des nombres n, l, m_l, m_s. L'application du principe de Pauli sur lequel se fonde la théorie du système périodique conduit à conclure que *dans un état quantique donné on ne peut trouver qu'un seul électron.* On n'a pas à individualiser les électrons puisqu'il importe de savoir seulement combien d'électrons se trouvent dans un état donné (n, l, m_l, m_s) et peu importe quels sont ces électrons. Si deux électrons échangent leurs coordonnées, l'état de l'enveloppe électronique *n'en sera absolument pas modifié.* Le nombre quantique de spin m_s peut être éliminé de la discussion puisqu'il ne peut prendre que deux valeurs $m_s = \pm 1/2$. Dans ce cas on pourra dire que *dans l'état quantique* (n, l, m_l) *ne peuvent se trouver plus de deux électrons.*

3. L'ensemble des électrons d'un atome ayant une valeur donnée du nombre quantique principal n forment une *couche électronique* ou *couche* tout court (ou *niveau*). Conformément à la terminologie en vigueur en spectroscopie des rayons X, les différentes couches sont dénotées par des lettres majuscules de l'alphabet latin:

n	1	2	3	4	5
Couche	K	L	M	N	O
Nombre maximal d'électrons dans la couche	2	8	18	32	50

L'ensemble des électrons ayant des valeurs données de n et l forment une *sous-couche* (*sous-niveau*). Les différents états d'une sous-couche diffèrent entre eux par la valeur du nombre quantique m_l. Comme ce dernier peut prendre les valeurs $m_l = -l, -(l-1), \ldots$ $\ldots, 0, \ldots +(l-1), +l$, une sous-couche de nombre quantique l peut comporter $2(2l+1)$ électrons. Les différentes sous-couches correspondant à différentes valeurs de l sont désignées par les lettres minuscules de l'alphabet latin:

l	0	1	2	3	4
Sous-couche	s	p	d	f	g
Nombre maximal d'électrons dans la sous-couche	2	6	10	14	18

Le symbole désignant les couches, les sous-couches et les nombres d'électrons contenus dans chaque sous-couche est appelé *configuration électronique* de l'atome. Par exemple, $1s^2 2s^2 2p^6$ signifie que dans l'état $n = 1$, $l = 0$ se trouvent deux électrons, dans l'état $n = 2$, $l = 0$ se trouvent deux autres électrons, et dans l'état $n =$ $= 2$, $l = 1$ il y en a six. On verra par la suite que ce symbole désigne la configuration électronique du néon. Le nombre maximal d'électrons contenus dans une couche à n donné s'obtient en additionnant les nombres d'électrons contenus dans toutes les sous-couches ayant des valeurs admissibles de l:

$$\sum_{l=0}^{n-1} 2(2l+1) = 2n^2. \qquad (47.1)$$

Les nombres correspondants pour différentes valeurs de n ont été indiqués plus haut.

4. Examinons maintenant comment varient les configurations électroniques lorsqu'on passe d'une espèce d'atome à une autre dans l'ordre de croissance de leurs numéros atomiques Z. Lorsque Z augmente d'une unité, la charge électrique du noyau augmente d'une unité et un électron vient s'ajouter à l'enveloppe électronique de l'atome. Le principe régissant la variation concomitante de la configuration électronique consiste en ce que la nouvelle configuration de $Z+1$ électrons doit avoir *la plus petite énergie possible* parmi

toutes celles que peut accepter la mécanique quantique. Autrement dit, l'énergie de liaison du nouvel électron associé à l'atome doit avoir *la plus grande valeur possible*. L'application réelle de ce principe exigerait la résolution du problème quantique de plusieurs corps, ce qui est pratiquement irréalisable. On en est réduit à utiliser des données empiriques chimiques et plus particulièrement des données spectroscopiques relatives aux potentiels d'ionisation des atomes. La théorie du système périodique acquiert de ce fait un caractère semi-empirique et plutôt descriptif. On ferait mieux de parler non pas de théorie, mais d'*interprétation* du système périodique comme nous l'avons fait dans le titre de ce paragraphe.

Il paraîtrait normal que ces couches se remplissent successivement les unes à la suite des autres et que dans une couche donnée se remplissent entièrement d'abord la sous-couche s, puis dans l'ordre les sous-couches l, d et f. En réalité, cet ordre « idéal » de remplissage est en désaccord avec le principe qui vient d'être énoncé. Les écarts à cet ordre sont pour l'essentiel dus à ce que les électrons possèdent des moments cinétiques orbitaux $L = \sqrt{l(l+1)}\hbar$. L'énergie de liaison de l'électron dépend non seulement de son énergie potentielle dans le champ électrique produit par le noyau et son cortège électronique, mais encore de l'énergie « centrifuge » $\frac{L^2}{2m_e r^2} = \frac{l(l+1)\hbar^2}{2m_e r^2}$, où r est la distance entre l'électron et le noyau. Les énergies « centrifuge » et électrique ont des signes contraires et exercent donc des actions en sens opposés. Pour faire image, on dira que la force centrifuge, qui apparaît lors de la rotation orbitale de l'électron, tend à éloigner l'électron du noyau. C'est pour cette raison que chacun des dix électrons de la sous-couche $3d$ possède une énergie de liaison plus petite que celle de chacun des deux électrons de la sous-couche $4s$. C'est pour cela que la sous-couche $4s$ se remplit d'électrons avant la sous-couche $3d$, quoique le nombre quantique principal soit dans le second cas plus petit que dans le premier. L'énergie « centrifuge » est particulièrement grande pour les sous-couches d et f, puisqu'on a alors respectivement $l(l+1) = 6$ et $l(l+1) = 12$. C'est la présence de ces sous-couches qui détermine les écarts à l'ordre « idéal » de remplissage dont il a été question plus haut.

En réalité les couches et les sous-couches, en faisant abstraction de certains détails indiqués dans le tableau 6, se remplissent d'électrons dans l'ordre suivant:

$1s^2$	2 électrons,
$2s^2 2p^6$	8 électrons,
$3s^2 3p^6$	8 électrons,
$4s^2 3d^{10} 4p^6$	18 électrons,
$5s^2 4d^{10} 5p^6$	18 électrons,
$6s^2 4f^{14} 5d^{10} 6p^6$	32 électrons,
$7s^2 5f^{14} 6d^{10} 7p^6$	32 électrons.

5. La couche $n = 1$ se confond avec la sous-couche s ($l = 0$). Dans l'atome d'hydrogène cette sous-couche (état $1s$) est occupée par un seul électron. Dans l'atome d'hélium un deuxième électron se trouvant dans l'état $1s$ se place sur cette sous-couche. Mais dans l'atome d'hélium l'énergie de liaison d'un électron est près de deux fois plus grande que l'énergie de liaison de l'électron de l'atome d'hydrogène. La raison en est que la charge du noyau d'hélium étant deux fois plus grande que celle du noyau d'hydrogène, l'électron de l'atome d'hélium à l'état normal se trouve plus près du noyau que l'électron dans l'atome d'hydrogène. D'autre part, la présence d'un second électron réduit l'énergie de liaison du premier électron. Ces deux éléments — hydrogène et hélium — forment la *première période* du système périodique.

Ajoutons maintenant à l'atome un troisième électron en augmentant simultanément la charge du noyau d'une unité. Le troisième électron ne peut se loger dans la couche K ($n = 1$), puisque cette couche est déjà complète. Ce troisième électron commence l'occupation de la sous-couche s de la couche L ($n = 2$) et prend l'état $2s$. On obtient ainsi le métal alcalin $_3$Li. Un quatrième électron vient se placer sur la sous-couche $2s$ et on obtient l'élément béryllium $_4$Be. Un cinquième électron ne pourra pas se loger dans la sous-couche $2s$ puisqu'elle est déjà remplie. Par conséquent, à partir de l'élément $_5$B commence à se remplir la sous-couche $2p$. L'occupation de cette couche par de nouveaux électrons fournit les éléments $_6$C, $_7$N, $_8$O, $_9$F. Le remplissage de la sous-couche $2p$ s'achève avec le gaz rare $_{10}$Ne. Ainsi s'est constituée la *deuxième période* (*courte*) comportant 8 éléments du système périodique.

A partir de l'élément alcalin $_{11}$Na commence l'occupation de la couche M ($n = 3$). Mais dès que les sous-couches s et p sont entièrement occupées, on obtient le gaz rare $_{18}$Ar. C'est la *troisième période*, également *courte*, puisqu'elle ne comporte que 8 éléments. A partir de là commencent les écarts à l'ordre de remplissage « idéal » des couches et des sous-couches. En effet, ici commence l'occupation de la sous-couche $4s$ et ce n'est que lorsque celle-ci aura été remplie que commencera l'occupation de la sous-couche $3d$ laissée provisoirement vacante et même l'occupation de cette sous-couche présente certaines irrégularités.

6. L'ordre d'occupation successive des couches et des sous-couches est indiqué dans le tableau 6 où sont répertoriées les configurations électroniques de tous les éléments. Tous les éléments sont partagés en deux groupes; les éléments du premier groupe se trouvent à gauche et ceux du second groupe sont décalés vers la droite. On procède de la même façon avec les symboles chimiques des éléments lorsqu'on dresse un tableau du système périodique permettant de disposer les uns au-dessus des autres les éléments ayant des propriétés chimiques analogues. Dans les atomes des éléments

Tableau 6

Configurations électroniques des éléments

Période	Elément		Configuration électronique	Potentiel d'ionisation, en V
1	1 H hydrogène		$1s^1$	13,539
	2 He hélium		$1s^2$	24,45
2	3 Li lithium	Configuration de l'hélium +	$2s^1$	5,37
	4 Be béryllium		$2s^2$	9,48
	5 B bore		$2s^2 2p^1$	8,4
	6 C carbone		$2s^2 2p^2$	11,217
	7 N azote		$2s^2 2p^3$	14,47
	8 O oxygène		$2s^2 2p^4$	13,56
	9 F fluor		$2s^2 2p^5$	18,6
	10 Ne néon		$2s^2 2p^6$	21,48
3	11 Na sodium	Configuration du néon +	$3s^1$	2,12
	12 Mg magnésium		$3s^2$	7,61
	13 Al aluminium		$3s^2 3p^1$	5,96
	14 Si silicium		$3s^2 3p^2$	7,39
	15 P phosphore		$3s^2 3p^3$	10,3
	16 S soufre		$3s^2 3p^4$	10,31
	17 Cl chlore		$3s^2 3p^5$	12,96
	18 Ar argon		$3s^2 3p^6$	15,69
4	19 K potassium	Configuration de l'argon +	$4s^1$	4,32
	20 Ca calcium		$4s^2$	6,09
	21 Sc scandium		$4s^2 3d^1$	6,57
	22 Ti titane		$4s^2 3d^2$	6,80
	23 V vanadium		$4s^2 3d^3$	6,76
	24 Cr chrome		$4s^1 3d^5$	6,74
	25 Mn manganèse		$4s^2 3d^5$	7,40
	éléments de transition: 26 Fe fer		$4s^2 3d^6$	7,83
	27 Co cobalt		$4s^2 3d^7$	7,81
	28 Ni nickel		$4s^2 3d^8$	7,61
	29 Cu cuivre		$4s^1 3d^{10}$	7,62
	30 Zn zinc		$4s^2 3d^{10}$	9,35
	31 Ga gallium		$4s^2 4p^1 3d^{10}$	5,27
	32 Ge germanium		$4s^2 4p^2 3d^{10}$	7,85
	33 As arsenic		$4s^2 4p^3 3d^{10}$	9,4
	34 Se sélénium		$4s^2 4p^4 3d^{10}$	
	35 Br brome		$4s^2 4p^5 3d^{10}$	11,8
	36 Kr krypton		$4s^2 4p^6 3d^{10}$	13,94

Suite

Période		Elément		Configuration électronique	Potentiel d'ionisation, en V
5		37 Rb rubidium	Configuration du crypton +	$5s^1$	4,16
		38 Sr strontium		$5s^2$	5,67
		39 Y yttrium		$5s^24d^1$	6,5
		40 Zr zirconium		$5s^24d^2$	
		41 Nb niobium		$5s^24d^4$	
		42 Mo molybdène		$5s^14d^5$	7,65
		43 Tc technétium		$5s^24d^5$	
	Eléments de transition	44 Ru ruthénium		$5s^14d^7$	7,7
		45 Rh rhodium		$5s^24d^8$	7,7
		46 Pd palladium		$4d^{10}$	8,5
		47 Ag argent		$5s^14d^{10}$	7,54
		48 Cd cadmium		$5s^24d^{10}$	8,95
		49 In indium		$5s^25p^14d^{10}$	5,76
		50 Sn étain		$5s^25p^24d^{10}$	7,37
		51 Sb antimoine		$5s^25p^34d^{10}$	8,5
		52 Te tellure		$5s^25p^44d^{10}$	
		53 I iode		$5s^25p^54d^{10}$	10,44
		54 Xe xénon		$5s^25p^64d^{10}$	12,08
6		55 Cs césium	Configuration du xénon +	$6s^2$	3,88
		56 Ba baryym		$6s^25d^1$	5,19
		57 La lanthane		$6s^24f^2$	
	Lanthanides	58 Ce cérium		$6s^24f^3$	
		59 Pr praséodyme		$6s^24f^4$	
		60 Nd néodyme		$6s^24f^5$	
		61 Pm prométhium		$6s^24f^6$	
		62 Sm samarium		$6s^24f^7$	
		63 Eu europium		$6s^25d^14f^7$	
		64 Gd gadolinium		$6s^25d^14f^8$	
		65 Tb terbium		$6s^24f^{10}$	
		66 Dy dysprosium		$6s^24f^{11}$	
		67 Ho holmium		$6s^24f^{12}$	
		68 Er erbium		$6s^24f^{13}$	
		69 Tu thulium		$6s^24f^{14}$	
		70 Yb ytterbium		$6s^25d^14f^{14}$	
		71 Lu lutétium			
		72 Hf hafnium		$6s^25d^24f^{14}$	
		73 Ta tantale		$6s^25d^34f^{14}$	
		74 W tungstène		$6s^25d^44f^{14}$	
		75 Re rhénium		$6s^25d^54f^{14}$	
	Eléments de transition	76 Os osmium		$6s^15d^74f^{14}$	
		77 Ir iridium		$6s^25d^74f^{14}$	
		78 Pt platine		$6s^15d^94f^{14}$	

Suite

Période	Élément			Configuration électronique	Potentiel d'ionisation, en V
6		79 Au or		$6s^15d^{10}4f^{14}$	
		80 Hg mercure		$6s^25d^{10}4f^{14}$	9,20
	81 Tl	thallium		$6s^26p^15d^{10}4f^{14}$	10,59
	82 Pb	plomb		$6s^26p^25d^{10}4f^{14}$	7,39
	83 Bi	bismuth		$6s^26p^35d^{10}4f^{14}$	8,0
	84 Po	polonium		$6s^26p^45d^{10}4f^{14}$	
	85 At	astate		$6s^26p^55d^{10}4f^{14}$	
	86 Rn	radon		$6s^26p^65d^{10}4f^{14}$	10,69
7	87 Fr	francium	Configuration du radon +	$7s^1$	
	88 Ra	radium		$7s^2$	
		89 Ac actinium		$7s^26d^1$	
	Actinides	90 Th thorium		$7s^26d^2$	
		91 Pa protactinium		$7s^26d^15f^2$	
		92 U uranium		$7s^26d^15f^3$	
		93 Np neptunium		$7s^26d^15f^4$	
		94 Pu plutonium		$7s^25f^6$	
		95 Am américium		$7s^25f^7$	
		96 Cm curium		$7s^26d^15f^7$	
		97 Bk berkélium		$7s^26d^15f^8$	
		98 Cf californium		$7s^25f^{10}$	
		99 Es einsteinium		$7s^25f^{11}$	
		100 Fm fermium		$7s^25f^{12}$	
		101 Md mendélévium		$7s^25f^{13}$	
		102 Jl joliotium		$7s^25f^{14}$	
		103 Rf rutherfordium		$7s^26d^15f^{14}$	
		104 (Ku courtchatovium)		$7s^26d^25f^{14}$	
		105 (Ns nielsbohrium)		$7s^26d^35f^{14}$	
		106		$7s^26d^45f^{14}$	
		107		$7s^26d^55f^{14}$	
		109			

du premier groupe, les sous-couches s et p sont occupées et leur remplissage s'effectue de façon régulière. C'est d'abord la sous-couche s qui doit être entièrement remplie, et ce n'est qu'ensuite que commence l'occupation de la sous-couche p. A la configuration électronique de chaque élément on ajoute un électron s ou un électron p et on obtient alors la configuration électronique de l'élément suivant.

Dans les atomes des éléments du deuxième groupe, les électrons vont se loger sur les sous-couches d et f, et cela constitue un écart à l'ordre d'occupation « idéal » dont il a été question plus haut. C'est ainsi que dans la *quatrième période*, à la suite de la configuration électronique du vanadium $4s^23d^3$ vient la configuration électronique du chrome $4s^13d^5$ (on a omis, pour simplifier, la configuration de l'argon), ce qui signifie que dans la sous-couche $3d$ le nombre

d'électrons a augmenté de deux unités, l'un de ces électrons provenant de la sous-couche $4s$. Dans la configuration électronique de l'élément suivant — manganèse $4s^2 3d^5$, le nouvel électron vient se loger dans la sous-couche $4s$ et non pas dans la sous-couche $3d$, ce qui rétablit le nombre d'électrons $4s$. Un examen du tableau 6 montre l'existence d'autres irrégularités de remplissage des sous-couches d et f.

Chaque période du système périodique débute par un métal alcalin, dans la configuration électronique duquel il n'y a qu'un seul électron s périphérique. De tous les éléments, ce sont les atomes des métaux alcalins qui ont les plus petits potentiels d'ionisation. C'est pour cela que ces atomes cèdent aisément leurs électrons s périphériques aux atomes d'autres éléments et c'est ce qui explique la grande activité chimique des métaux alcalins.

Chaque période s'achève par les gaz rares He, Ne, Ar, Kr, Xe, Rn. La sous-couche s-p externe des atomes des gaz rares (He excepté) contient 8 électrons qui constituent un système particulièrement compact, symétrique et fortement lié. De ce fait les atomes des gaz rares présentent les plus hauts potentiels d'ionisation. De là vient leur manque d'activité chimique avec d'autres éléments. Par contre, les éléments du groupe voisin, le septième, du système périodique — fluor, chlore, brome, iode, présentent une très grande activité chimique. Cela tient à ce que leurs sous-couches s-p externes ne contiennent chacune que sept électrons et qu'il est facile de les compléter jusqu'à l'octet par capture d'un huitième électron.

Chaque demi-période s'achève par des éléments dits *de transition*; dans le tableau 6 ces éléments sont encadrés en pointillé. Ils sont groupés en triades : (fer, cobalt, nickel), (ruthénium, rhodium, palladium), (osmium, iridium, platine); ces triades sont généralement placées dans une seule case du système périodique.

On doit accorder une attention particulière à 14 éléments du cérium au lutétium y compris, que l'on appelle *terres rares* ou *lanthanides*. Dans ces éléments se réalise l'occupation progressive de la sous-couche interne $4f$. Comme les sous-couches externes restent pratiquement inchangées, tous les éléments des terres rares possèdent des propriétés chimiques très semblables et il est par suite difficile de les séparer. Dans le tableau 6 les lanthanides sont encadrés en trait plein.

L'histoire de la découverte de l'élément 72 (hafnium) est fort instructive. Jusqu'en 1922, cet élément était inconnu, mais une place lui était réservée de façon erronée parmi les terres rares. Mais Bohr signala que selon des considérations théoriques le groupe des terres rares devait s'arrêter au 71-ième élément (lutétium) et que l'élément 72 devrait présenter des propriétés chimiques analogues à celles du zirconium ($Z = 40$). On se mit alors à le chercher et on le décela effectivement dans les minerais de zirconium.

Les *actinides* se comportent à la façon des lanthanides. On désigne sous ce nom les 14 éléments allant du thorium au rutherfordium dont la plupart ont été préparés artificiellement. Dans le tableau 6 ces éléments sont également encadrés en trait plein. Dans les atomes de ces éléments s'effectue l'occupation progressive de la sous-couche interne $5f$, les sous-couches externes restant pratiquement inchangées. Aussi, à l'instar des lanthanides, les actinides ont tous des propriétés chimiques peu différentes.

Le système périodique s'arrête à l'élément n° 109, le dernier qui a été obtenu artificiellement. Cela ne signifie nullement qu'on n'arrivera pas à produire plus tard d'autres éléments de plus grands numéros atomiques.

7. Les acquis de la théorie du système périodique des éléments chimiques sont indubitables, mais on ne peut passer sous silence ses défauts de principe. On a fait état de l'un de ses défauts au sous-paragraphe 2, il nous reste à en noter un autre encore. Les électrons appartenant aux différentes sous-couches se caractérisent par des nombres quantiques orbitaux l. Cela implique la conservation de la valeur numérique du moment cinétique orbital de chaque électron. Or la loi de la conservation du moment cinétique n'est valable que pour des particules se déplaçant dans un champ de force *à symétrie centrale* et n'est pas vérifiée dans d'autres cas. Le champ électrique dans lequel se meut un électron donn dans l'atome est créé par le noyau et tous les autres électrons, et il n'a pas une symétrie sphérique. Même dans le cas où il n'y a qu'un seul électron périphérique, toutes les couches internes étant entièrement remplies d'électrons, la symétrie sphérique s'obtient grâce à un moyennage quantique. Or dans l'équation de Schrödinger qui décrit le mouvement des électrons, figure non pas l'énergie potentielle moyenne, mais l'énergie potentielle définie classiquement en tant que fonction des coordonnées de tous les électrons ponctuels (cf. § 21, 3).

§ 48. Les rayons X

1. Les rayons X produits dans les tubes sont émis lorsqu'un faisceau d'électrons rapides est arrêté par l'anode (dans les tubes de construction ancienne sur l'anticathode, cf. t. III, § 117). Les propriétés ondulatoires des rayons X ont été démontrées en 1912 par Laue (1879-1960) et ses collaborateurs Friedrich (1883-1968) et Knipping (1883-1935) qui ont réalisé leur diffraction par un cristal (cf. t. IV, § 61). Dès 1905, Barkla (1877-1944) a montré que si les rayons X étaient effectivement des ondes, ces ondes devraient être *transversales*. Dans les expériences de Barkla un faisceau S de rayons X était diffusé par un corps A (fig. 86). On décelait la diffusion des rayons X à l'aide d'une chambre d'ionisation; en y pénétrant, les rayons X provoquaient l'ionisation de l'air qui, de ce fait, devenait

conducteur. Le rayonnement diffusé sous un angle de 90° tombait sur un corps *B* et était diffusé une nouvelle fois. Il s'avéra que l'intensité de ce dernier rayonnement diffusé était nulle suivant la direction *BC* orthogonale au plan *SAB*, tandis que dans toutes les autres directions, y compris la direction *BD* antiparallèle à *SA*, la diffusion avait bien lieu. Barkla en conclut que les ondes X, si les rayons X étaient bien des ondes, devaient être transversales.

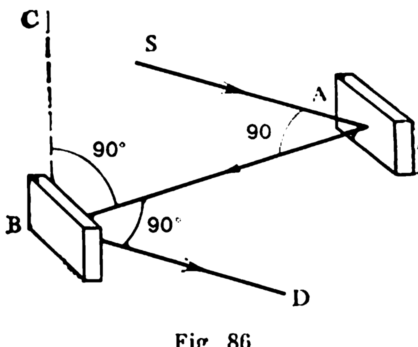


Fig. 86

En effet, si les ondes sont transversales, le vecteur électrique du rayon incident *SA*, ainsi que les vibrations des électrons excités dans le corps *A* peuvent être décomposés suivant les directions *AB* et *BC*. Les vibrations suivant *AB* ne produisent aucun rayonnement puisqu'un électron vibrant ne rayonne pas suivant la direction de ses vibrations. Seule l'onde dont le vecteur électrique est parallèle à *BC* parvient jusqu'au corps *B*. Cette onde excitera dans le corps *B* des vibrations électroniques suivant la direction *BC*. Par conséquent, il n'y aura aucun rayonnement suivant cette direction.

On conçoit que dans les expériences de Barkla les rayons X devaient être assez « durs » pour ne pas être absorbés notablement pendant leur trajet dans l'air. Pour cela la tension appliquée au tube devait être suffisamment élevée. Pour que l'absorption dans le corps diffusant soit assez faible, les corps *A* et *B* devaient être faits en matériaux de faible numéro atomique *Z* (charbon dans les expériences de Barkla). S'il avait utilisé des substances de grand *Z*, avec les hautes tensions appliquées au tube, il aurait obtenu un rayonnement de raies de grande intensité caractéristique du matériau. Le charbon, la paraffine, etc., ne produisent qu'un rayonnement caractéristique faible et mou, qui est absorbé par l'air sur un trajet de quelques centimètres seulement et qui ne nuit donc nullement aux expériences.

2. C'est encore Barkla qui constata que les rayons X issus de l'anticathode se composaient de deux parties. La première partie résulte de la *décélération* (Bremsstrahlung) des électrons dans l'anticathode. Cette partie du rayonnement ne dépend absolument pas du matériau de l'anticathode. L'autre partie du rayonnement dépend essentiellement du matériau dont est faite l'anticathode. Barkla caractérisait le rayonnement X par sa *dureté*, i.e. son aptitude à traverser différentes substances. Ce n'est qu'après les expériences de Laue, lorsque furent élaborées les méthodes de radioscopie et de

mesure des longueurs d'onde des rayons X, qu'on a pu donner une caractéristique quantitative précise des rayons X (cf. t. IV, § 61).

Le *rayonnement de freinage* a un *spectre continu*. Par analogie avec la lumière blanche, on l'appelle parfois *rayonnement* X *blanc*. Du côté des grandes longueurs d'onde l'intensité du rayonnement de freinage décroît lentement et tend asymptotiquement vers zéro. Du côté des ondes courtes le spectre continu est limité par une frontière assez abrupte (cf. fig. 5). On a déjà expliqué au § 2 que ces particularités du spectre continu de rayons X étaient liées à leur *nature quantique*. La frontière côté ondes courtes du spectre continu qui est déterminée par la formule (2.6), ne dépend absolument pas du matériau de l'anticathode et sa position est fonction de la tension appliquée aux tubes. D'une façon générale, la répartition relative de l'énergie suivant les longueurs d'onde dans le spectre du rayonnement X de freinage ne dépend pas du matériau de l'anticathode. La nature de l'anticathode n'exerce son influence que sur l'intensité intégrale du rayonnement.

Le *rayonnement caractéristique* se présente sous la forme d'un *spectre de raies*, i.e. se compose d'une suite régulière de raies spectrales assez fines. Les longueurs d'onde de ces raies dépendent uniquement de la nature de l'anode. A ce point de vue le rayonnement X caractéristique rappelle le spectre de raies des gaz dans la région visible. Le bombardement électronique de l'anode produit simultanément l'émission d'un spectre continu et d'un spectre de raies, tandis que le bombardement par des particules α ou par des protons ne donne lieu qu'à l'émission d'un spectre de raies. Les rayons X qui sont émis par une substance lorsqu'on l'irradie avec d'autres rayons X comportent pour une part un faisceau primaire diffusé et pour une part le rayonnement X caractéristique propre de cette substance. Le spectre de raies apparaît lorsque la tension appliquée au tube dépasse une certaine valeur liée uniquement à la nature de l'anode.

La figure 87 représente des courbes expérimentales de la répartition spectrale de l'énergie des radiations X émises par des anticathodes en tungstène, en molybdène et en chrome, lorsque la tension appliquée au tube est $V = 35\,000$ V. L'intensité du rayonnement est mesurée par le courant traversant la chambre d'ionisation. Avec une anode en molybdène on observe, en superposition au spectre continu, deux raies caractéristiques de longueurs d'onde $\lambda = 0{,}063$ et $\lambda = 0{,}071$ nm. Le rayonnement caractéristique apparaît aussi dans le cas d'une anticathode en chrome, mais à de plus grandes longueurs d'onde, sortant du cadre de la figure. Dans le cas du tungstène, la tension appliquée au tube $V = 35\,000$ V est insuffisante pour exciter des raies caractéristiques dans la gamme des longueurs d'onde utilisées sur la figure et on ne décèle que le spectre continu.

3. Chaque raie du rayonnement X caractéristique est liée au passage d'un atome d'un niveau énergétique initial sur un autre niveau. La fréquence de la raie est donnée par la règle des fréquences de Bohr

$$\hbar\omega = \mathscr{E}_1 - \mathscr{E}_2. \qquad (48.1)$$

Comme la longueur d'onde des rayons X est petite, la différence d'énergie entre l'état initial et l'état final de l'atome est très grande;

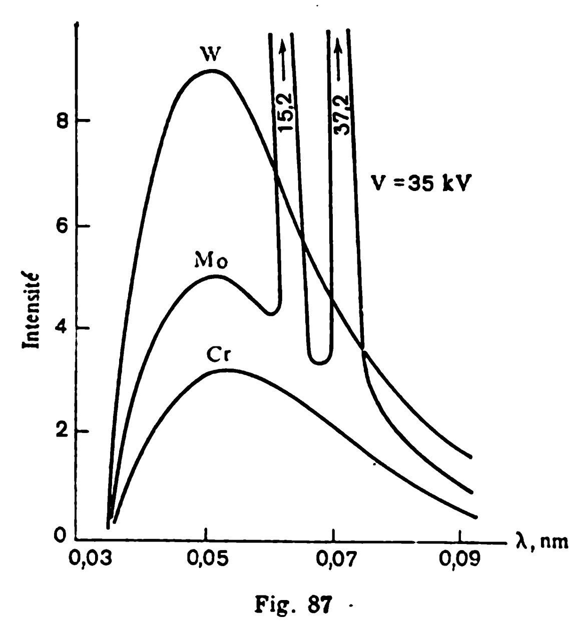


Fig. 87

pour les éléments lourds, cette différence est de 10^3 à 10^4 et même 10^5 fois plus grande que les différences auxquelles on a affaire dans la gamme optique. On en conclut que les transitions quantiques concernent les électrons des couches *internes* et non les électrons de valence des atomes. Or pour que ces transitions soient possibles, il faut qu'il y ait dans les profondeurs de l'enveloppe électronique des *places vacantes*, non occupées par des électrons où viendraient se placer les électrons provenant d'autres états quantiques de l'atome. Ces places vacantes sont créées en bombardant les atomes avec des électrons rapides, des photons de haute énergie ou avec d'autres particules rapides.

Supposons, par exemple, qu'on extrait un électron de la couche profonde *K* d'un atome. L'énergie de l'atome s'en trouvera accrue de la valeur de l'énergie d'ionisation de la couche *K*. Soit $\mathscr{E}_1$ l'énergie que possède l'atome après extraction d'un électron de la couche *K*. La place laissée vacante par cet électron sera occupée par un élec-

tron de la couche L. La couche K sera remplie, tandis qu'un électron manquera à la couche L. On a maintenant un atome qui a perdu un électron de la couche L. Son énergie $\mathscr{E}_2$ est plus petite que l'énergie $\mathscr{E}_1$. Si cette transition d'un électron de la couche L sur la couche K donne lieu à l'émission d'un quantum X, sa fréquence sera donnée par la formule (48.1). La place laissée vacante dans la couche L sera occupée par un électron provenant de la couche M et cette transition s'accompagnera de l'émission d'un quantum X de fréquence plus basse. Il en sera de même pour la transition d'un électron de la couche N, etc. Il va de soi que les transitions ne se réalisent pas uniquement entre couches voisines, mais aussi entre les couches M et K, N et K, etc. Ces dernières transitions s'accompagnent de l'émission de radiations plus dures que dans le cas de la transition $L \rightarrow K$. L'ensemble de tous ces processus conduit à l'apparition de la *totalité du spectre* de raies du rayonnement X caractéristique de l'atome considéré.

La fréquence des raies du rayonnement caractéristique dépend de la différence des énergies des niveaux énergétiques de l'atome qui est calculée en supposant qu'*un seul électron a été extrait* de chacun de ces niveaux. Dans ce cas l'atome passe à un état excité et son énergie s'accroît de la valeur de l'énergie d'ionisation ; cette dernière est d'autant plus grande que le niveau concerné de l'atome non excité est plus profond. Ces considérations montrent pourquoi on utilise, pour expliquer la production du rayonnement X, le schéma des niveaux d'un atome excité, schéma établi en supposant qu'un seul électron a été enlevé du niveau considéré, le nombre d'électrons occupant tous les autres niveaux étant inchangé. Il résulte de ce qui précède que le schéma des niveaux d'un atome ainsi excité est inversé par rapport au schéma des niveaux énergétiques de l'atome non excité. Cela signifie que dans ce schéma énergétique inversé un niveau énergétique est d'autant plus haut placé dans l'échelle des énergies qu'il était plus profond dans l'atome non excité (avant qu'un électron ait été enlevé de ce niveau).

La figure 88 représente un schéma inversé des niveaux énergétiques d'un atome. On sait que dans un atome non excité les couches électroniques correspondant aux nombres quantiques principaux $n = 1, 2, 3, 4, \ldots$ sont désignées par $K, L, M, \ldots$ On utilise les mêmes notations dans le cas de l'atome excité de la façon indiquée ci-dessus. On applique aussi à l'atome excité les notations spectroscopiques telles que $2^2S_{1/2}$, $2^2P_{1/2}$, . . . L'utilisation de ces notations est justifiée par ce que les sous-couches électroniques internes étant complètes, leur moment cinétique est nul ; lorsqu'un électron est éloigné, le moment cinétique qu'il emporte est transmis avec le signe contraire à la sous-couche d'origine de l'électron.

En champ magnétique nul l'énergie ne dépend pas du nombre quantique magnétique et ne dépend pratiquement que des nombres

quantiques n et j. La couche K se réduit à un seul niveau énergétique $1^2S_{1/2}$ ($n = 1$, $l = 0$, $j = 1/2$). La couche L ($n = 2$) se partage en trois sous-couches dénotées par L_I, L_{II}, L_{III}. La sous-couche L_I comporte un niveau $2S$. Les deux autres sous-couches $2P$ comportent

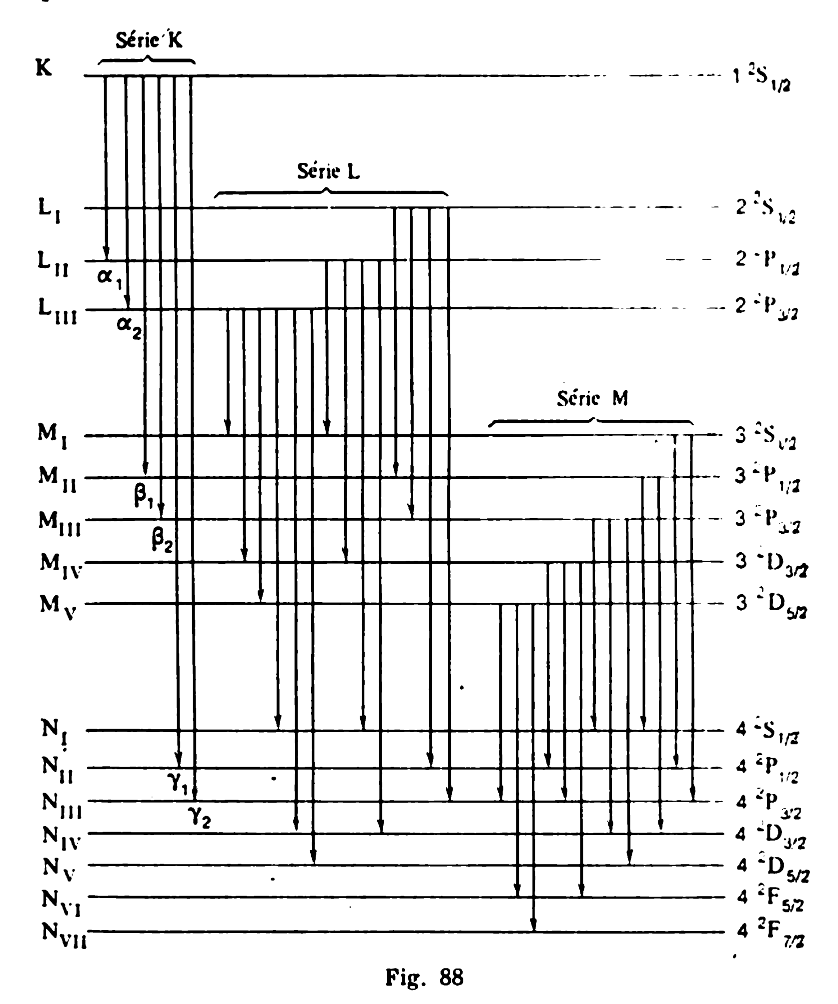


Fig. 88

les sous-niveaux $2^2P_{1/2}$ et $2^2P_{3/2}$ dont les énergies sont différentes en raison de l'interaction spin-orbite. La couche M ($n = 3$) comporte cinq sous-couches, la couche N ($n = 4$) en comporte sept (voir fig. 88), etc.

4. On peut considérer le manque d'un électron dans une sous-couche complète d'un atome comme un « trou ». Tout formellement

on peut poser que le « trou » peut se trouver dans différents états quantiques et passer d'un état à un autre. Lorsqu'une sous-couche comporte un trou, elle est incomplète, mais redevient complète lorsque le trou migre vers une autre sous-couche et c'est cette dernière qui devient incomplète. En se plaçant à ce point de vue, le schéma inversé des niveaux énergétiques profonds de l'atome est analogue au schéma usuel des niveaux d'énergie d'un atome n'ayant qu'un seul électron de valence. Dans ces conditions toutes les règles de sélection relatives à l'émission restent également en vigueur lorsque l'atome (le trou) passe d'un état quantique à un autre. Les règles de sélection sont

$$\Delta J = -1,\ 0,\ +1;\quad \Delta L = \pm 1. \tag{48.2}$$

Toutes les transitions satisfaisant à ces règles de sélection sont représentées par des flèches sur la figure 88. Ces flèches indiquent les transitions d'un atome excité depuis des niveaux de plus grande énergie à des niveaux d'énergie plus faible, ou ce qui revient au même, les transitions d'un trou d'un niveau à un autre. On n'a envisagé que les couches K, L, M et N. Il est facile de compléter la figure 88 en envisageant, si besoin est, des transitions sur les sous-niveaux des couches O ($n = 5$) et P ($n = 6$).

Le spectre de raies du rayonnement X se compose de séries de raies que l'on désigne par les lettres K, L, M, N, O. La série K résulte des transitions de l'atome excité du niveau K sur les sous-niveaux d'énergie plus petite des couches L, M, N, . . . La série L résulte de transitions analogues ayant pour point de départ les sous-niveaux de la couche L, la série M a pour origine les transitions depuis les sous-niveaux de la couche M, etc. On peut constater sur la figure que les raies de la série K sont des doublets. Les composantes des doublets sont notées respectivement α_1, α_2; β_1, β_2; γ_1, γ_2 (fig. 88). Les séries L, M, N ont une structure en multiplet plus compliquée. Les raies appartenant à ces multiplets sont désignées par des lettres grecques affectées d'indices numériques. La lettre α indique que la transition se produit entre des couches adjacentes, la lettre β que la transition s'effectue par-dessus une couche intermédiaire, etc. Les indices numériques dont sont affectées les lettres α, β, γ servent à numéroter les raies dans l'ordre des longueurs d'onde décroissantes.

Les explications relatives à l'origine du spectre de raies permettent de conclure que l'excitation de la couche K la plus profonde fait apparaître non seulement la série K mais le *spectre des rayons* X *tout entier*. D'une façon générale, chaque fois qu'une série ou une certaine raie du rayonnement X se trouve excitée, on voit apparaître toutes les séries et toutes les autres raies de plus grandes longueurs d'onde.

Lorsqu'un électron est extrait d'une sous-couche interne hors de l'atome, il peut posséder n'importe quelle énergie cinétique. Dans ce cas son énergie à l'état final n'est plus quantifiée. Lors du passage de cet état non quantifié sur une place vacante d'une couche de l'atome, on obtient un *spectre* X *continu.*

5. A la différence des spectres de raies optiques qui sont compliqués et très variés, les spectres de raies du rayonnement X caractéristiques des différents éléments sont simples et semblables. Cela tient à ce qu'en passant d'un élément chimique au suivant, la structure des couches internes des atomes ne varie que fort peu. Lorsque le nombre de charge Z augmente d'une unité, les spectres de raies des rayons X des éléments concernés sont semblables et on ne constate qu'un léger déplacement de toutes les raies X dans le sens des courtes longueurs d'onde. Cette particularité des spectres des rayons X a été mise en évidence en 1913 par Moseley (1887-1915) qui l'interpréta dans le cadre de la théorie de Bohr. Moseley étudia de façon systématique les séries K et L du rayonnement X de 38 éléments. Il utilisait un spectrographe à cristal fonctionnant selon le principe des réflexions de Bragg (cf. t. IV, § 61) dans lequel la chambre d'ionisation était remplacée par une plaque photographique.

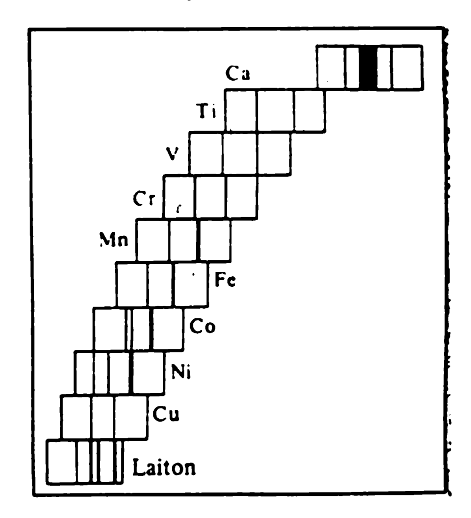


Fig. 89

La figure 89 reproduit la photographie des séries K des différents éléments obtenue par Moseley. La disposition relative des spectres des différents éléments est telle que la distance de n'importe quelle raie à l'extrémité gauche de la figure est approximativement proportionnelle à la longueur d'onde de cette raie. Les éléments sont disposés dans l'ordre de la croissance du numéro atomique en commençant par le calcium ($Z = 20$) jusqu'au zinc ($Z = 30$) entrant dans la composition du laiton. On observe une remarquable régularité de la croissance de la fréquence du rayonnement K caractéristique avec le numéro atomique de l'élément. La même régularité s'observe dans la croissance de la fréquence des raies L, M, N du rayonnement caractéristique en fonction du numéro atomique des éléments. La figure 89 montre clairement qu'entre le calcium et le titane devrait se trouver un élément de numéro atomique $Z = 21$.

Cet élément est le scandium dont l'existence a été prévue par Mendéléev et qui fut découvert en 1879 par Nilson et Cleve.

Les recherches de Moseley ont démontré par l'expérience que la grandeur fondamentale dont dépend la place d'un élément dans le système périodique est son *numéro atomique* et non pas sa masse atomique. D'autre part, les spectres de raies du rayonnement X caractéristique permettent de déterminer de façon univoque les numéros atomiques des éléments et de décider ainsi si toutes les cases du tableau périodique sont occupées ou s'il manque encore des éléments qu'il faudrait découvrir. Moseley réserva la case n° 43 pour un élément qui était inconnu à l'époque et qui, plus tard, fut obtenu artificiellement et appelé technétium. Avant les recherches de Moseley on ne savait pas lequel des éléments, le cobalt de masse atomique 58,933 ou le nickel de masse atomique 58,71, devait précéder l'autre dans le tableau périodique. Un examen de la figure 89 montre nettement que le cobalt doit être placé entre le fer et le nickel, quoique ta masse atomique soit plus grande que celle du nickel. La partie du système périodique comprise entre les cases $Z = 58$ et $Z = 71$ incluse est occupée par les éléments des terres rares. Leurs propriétés chimiques sont tellement proches et leurs masses atomiques étaient tellement mal connues qu'on avait des doutes sur leur ordre de succession dans le système périodique. L'étude des spectres de raies du rayonnement X et l'utilisation de la loi de Moseley (voir ci-dessous) permirent d'y rétablir l'ordre. Cette étude montra qu'entre l'hydrogène et l'uranium il doit se trouver exactement 92 éléments.

De la discussion sur l'origine du rayonnement X caractéristique, il ressort que la longueur d'onde des radiations émises dépend uniquement de la structure interne des couches électroniques des atomes. Autrement dit ce rayonnement est une *propriété atomique* de chaque élément chimique. De ce fait chaque élément entrant dans la composition d'un composé chimique ou d'un alliage donnera le même spectre caractéristique que s'il était seul. On s'en rend bien compte en examinant la figure 89. Le spectre du laiton résulte d'une simple superposition des spectres du cuivre et du zinc ; l'examen du spectre du cobalt montre que ce métal contient comme impuretés du fer et du nickel.

6. Moseley établit expérimentalement que la variation de la racine carrée de la fréquence des vibrations $\sqrt{\nu}$ d'une raie donnée de la série K en fonction du numéro atomique Z des éléments est traduite par une courbe régulière très proche de la ligne droite :

$$\sqrt{\nu} = M(Z-a), \qquad (48.3)$$

où M et a sont des constantes. La même formule mais avec des valeurs numériques différentes de M et a s'applique aux raies des sé-

ries L, M, N, O. Cette formule empirique exprime la *loi de Moseley* et détermine les déplacements des raies caractéristiques des rayons X lorsqu'on passe d'un élément chimique à un autre. Les mesures ultérieures plus précises ont mis en évidence certains écarts à la loi linéaire (48.3). Pour les raies des séries K et L ces écarts sont peu notables; pour les séries M, N et O les écarts peuvent être appréciables.

Il est impossible de démontrer ou d'indiquer la limite de précision de la loi de Moseley puisque cela implique le problème de plusieurs corps. On ne peut que donner une interprétation grossière de cette loi en ramenant le problème de plusieurs corps à un problème monoélectronique. Les résultats que l'on obtient ainsi sont pleinement satisfaisants, notamment pour les séries K et L, et permettent de dégager la signification physique des constantes figurant dans la formule (48.3).

Posons qu'un trou est apparu dans la couche K ne comportant que deux électrons, ce qui suppose qu'un électron a été éloigné hors de l'atome. Supposons que dans ce trou vient se loger un électron de la couche L et que ce processus s'accompagne de l'émission d'une raie K_α. Etudions le comportement ultérieur de cet électron jusqu'au moment où le trou dans la couche K sera comblé. On peut admettre que l'électron considéré ne subit aucune action de la part des électrons périphériques, ceux-ci constituant une enveloppe extérieure à charge à répartition sphérique. Le champ électrique agissant sur notre électron est produit par la charge du noyau et par l'électron restant dans la couche K. On peut remplacer ces deux particules par un noyau portant une charge « efficace » $(Z - 1)e$; tout se passe comme si l'unique électron resté dans la couche K faisait office d'écran vis-à-vis du champ du noyau.

Un raisonnement semblable, quoique moins concret, s'applique au cas où le trou est formé dans la couche L, trou qui sera occupé par la suite par un électron provenant de la couche M, le processus s'accompagnant de l'émission d'une raie L_α. Dans ce cas il convient d'étudier le comportement de l'électron de la couche M jusqu'au moment où il viendra se loger dans le trou de la couche L. L'effet d'écran est produit maintenant par les deux électrons de la couche K et les sept électrons de la couche L. Néanmoins, ultérieurement tant que l'électron considéré n'a pas rempli le trou de la couche L, certains électrons de la couche L cesseront de contribuer à l'effet d'écran. On posera en approximation que l'effet d'écran ne serait exercé que par la moitié environ des sept électrons restants de la couche L. Cela revient à poser que la charge efficace du noyau est égale à $(Z - a)e$, où a est la *constante d'écranage*. On peut s'attendre à ce que pour la raie L_α $a \approx 5{,}5$, valeur qui cadre à peu près avec les résultats de l'expérience, quoique cette estimation ne soit qu'une approximation grossière puisqu'il n'a pas été tenu compte, notamment,

de ce que la couche L se compose de trois sous-couches. La détermination de la valeur exacte des constantes d'écranage doit reposer sur l'expérience.

Dans les deux cas envisagés on a à résoudre un problème à un électron, comme ce fut le cas lorsqu'on avait cherché une interprétation des séries spectrales des métaux alcalins. Les fréquences des raies émises seront données par la formule

$$\nu = cR\,(Z-a)^2 \left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right), \tag{48.4}$$

où R est la constante de Rydberg (cf. § 13). C'est à partir de là qu'on obtient la formule (48.3) puisque pour les raies considérées les nombres quantiques n_1 et n_2 sont donnés. Pour la raie K_α $n_1 = 1$, $n_2 = 2$ et $a = 1$, comme on l'a établi ci-dessus. Par suite

$$\nu_{K_\alpha} = \frac{3}{4}\,cR\,(Z-1)^2. \tag{48.5}$$

Pour la raie L_α $n_1 = 2$, $n_2 = 3$, de sorte que

$$\nu_{L_\alpha} = \frac{5}{16}\,cR\,(Z-a_{L_\alpha})^2 \tag{48.6}$$

avec $a_{L_\alpha} \approx 5{,}5$.

7. Les spectres d'absorption des rayons X se caractérisent par ce qu'ils ne comportent aucune *inversion de raies*. Si on fait passer au travers d'une couche d'un élément quelconque un pinceau de rayons X obtenus par freinage (ce rayonnement forme un spectre continu), on n'observera pas à la place des raies caractéristiques des raies noires comme ce serait dans le cas des spectres d'absorption optiques. Cela tient à ce qu'*il n'y a pas d'atomes excités* dans la substance traversée par les rayons X. Toutes les couches électroniques internes étant entièrement remplies, il n'y a pas de « trous » où pourraient venir se loger, lors d'une irradiation par rayons X, des électrons provenant de niveaux énergétiques plus profonds.

L'absorption des rayons X par les substances est complètement indépendante de leurs propriétés optiques. Par exemple, un verre au plomb transparent, d'une épaisseur de quelques millimètres, absorbe pratiquement tous les rayons X incidents; c'est pour cela qu'on l'utilise pour protéger le personnel contre les irradiations par rayons X. Par contre, une feuille d'aluminium absolument opaque à la lumière visible, se laisse aisément traverser par les rayons X produits par les tubes à rayons X industriels (tension $\sim$100 kV); les feuilles d'aluminium sont transparentes aux rayons X.

8. Lorsqu'un pinceau parallèle de rayons X traverse une substance, il se trouve affaibli. Cet affaiblissement résulte de deux causes. Premièrement, de la *diffusion*: une partie des rayons subit une déflexion et se trouve ainsi exclue du pinceau parallèle; c'est un phénomène analogue à la diffusion de la lumière visible dans les mi-

lieux troubles. Dans ce dernier cas, le milieu est trouble soit parce qu'il contient en suspension des particules ayant un indice de réfraction différent, soit à cause des fluctuations de densité de la substance pure (ne contenant aucune particule étrangère) (cf. t. IV, § 98). La diffusion due aux fluctuations de densité est appelée *diffusion moléculaire* ou *diffusion de Rayleigh*. En raison de la petitesse de la longueur d'onde des rayons X, tous les milieux sont *troubles* vis-à-vis des rayons X. Dans ce cas, le rôle des inhomogénéités du milieu, déterminant la diffusion, est assumé par les atomes individuels ou même par les électrons des atomes et non par de petits fragments macroscopiques de la substance. Pour un rayonnement X mou dont les longueurs d'onde sont comparables aux dimensions atomiques, les électrons des atomes se comportent comme s'ils étaient liés et l'atome diffuse les rayons X comme un tout; de ce fait, les ondes diffusées par les différents électrons d'un atome sont *cohérentes*. Dans le cas d'un rayonnement X très dur, la longueur d'onde étant petite devant les dimensions de l'atome, il se produit une diffusion de Compton, *non cohérente*, des rayons X par les électrons des atomes.

On peut illustrer la différence quantitative existant entre la diffusion de la lumière visible et la diffusion des rayons X par l'exemple suivant. Par suite de la diffusion moléculaire un pinceau cylindrique de lumière visible s'affaiblit de e fois en traversant une couche de ≈ 1 km d'épaisseur d'une eau absolument pure, tandis que les rayons X y subissent le même affaiblissement relatif après traversée d'une épaisseur de ≈ 5 cm.

La dernière cause d'affaiblissement des rayons X résulte d'une *absorption vraie* lorsqu'une partie de l'énergie du pinceau de rayons X est finalement convertie en chaleur.

Si le pinceau cylindrique de rayons X est monochromatique (λ = const), l'affaiblissement de son intensité I après parcours d'une distance dx dans une substance homogène est donné par la relation $dI = -\mu I dx$, où la constante μ est appelée *coefficient d'affaiblissement* des rayons X. Par conséquent,

$$I(x) = I_0 e^{-\mu x}, \tag{48.7}$$

i.e. l'intensité diminue de e fois après le parcours d'une distance $x = \frac{1}{\mu}$. Comm el'affaiblissement de l'intensité d'un pinceau de rayons X résulte aussi bien d'une absorption vraie que d'une diffusion, on peut écrire

$$\mu = \tau + \sigma, \tag{48.8}$$

où τ est le *coefficient d'absorption vraie* et σ le *coefficient de diffusion* des rayons X.

Comme tous les coefficients μ, σ, τ sont proportionnels à la masse volumique ρ de la substance, il est préférable d'utiliser les *coeffi-*

cients massiques, i.e.

$$\mu_m = \frac{\mu}{\rho}, \quad \tau_m = \frac{\tau}{\rho}, \quad \sigma_m = \frac{\sigma}{\rho}. \tag{48.9}$$

En introduisant μ_m dans la formule (48.7), cette dernière s'écrit

$$I(x) = I_0 e^{-\mu_m \rho x}. \tag{48.10}$$

En notant S l'aire de la section droite du faisceau, la quantité $S\rho x$ représente la masse de substance traversée par le faisceau de rayons X sur une distance x. Il s'ensuit que ρx est la masse de la substance parcourue par le faisceau sur une distance x, si l'aire de la section droite du faisceau est égale à l'unité.

Dans les calculs théoriques il est encore plus commode d'utiliser les *coefficients* dits *atomiques*:

$$\mu_a = \frac{\mu}{\rho} m = \frac{\mu}{\rho} \frac{A}{N_{Av}}, \quad \tau_a = \frac{\tau}{\rho} m = \frac{\tau}{\rho} \frac{A}{N_{Av}},$$
$$\sigma_a = \frac{\sigma}{\rho} m = \frac{\sigma}{\rho} \frac{A}{N_{Av}}, \tag{48.11}$$

où m est la masse d'un atome égale à $m = A/N$, A la masse atomique et N le nombre d'Avogadro. Etant homogènes à une surface, ces coefficients peuvent être assimilés à la *section droite de l'atome* lorsqu'il s'agit de l'affaiblissement, de l'absorption ou de la diffusion des rayons X. Par exemple, μ_a caractérise l'affaiblissement d'intensité du faisceau de rayons X dans une couche contenant un atome par unité d'aire de la section droite du faisceau. Si dans les formules (48.11) on entend par A la masse moléculaire, les quantités définies par ces formules doivent être appelées *coefficients moléculaires*.

9. Généralement, le coefficient d'absorption des rayons X croît avec la longueur d'onde, mais pour certaines valeurs de celle-ci le coefficient d'absorption diminue brusquement, puis recommence à augmenter lentement. Cette particularité de l'absorption des rayons X trouve une explication physique simple. Supposons que la longueur d'onde des rayons X soit tellement petite que tous les niveaux profonds, y compris le niveau K, soient excités. L'énergie des rayons X incidents est utilisée à cette excitation. A mesure que la longueur d'onde augmente, à partir d'une certaine valeur λ, l'énergie du quantum X ne suffit plus pour exciter le niveau K, l'absorption s'en trouve brusquement diminuée et une *arête d'absorption* apparaît. A mesure que la longueur d'onde continue à augmenter, les niveaux L_I, L_{II}, L_{III} cessent successivement de s'exciter (pour un schéma des niveaux voir fig. 88), ce qui fait apparaître de nouvelles arêtes d'absorption, etc.

La relation empirique

$$\tau_a = CZ^4\lambda^3 \tag{48.12}$$

est vérifiée avec une bonne précision; la constante C diminue par saut chaque fois qu'on passe par l'arête d'absorption. Il résulte tout naturellement de (48.12) que

$$\tau_m = \frac{\tau_a}{\rho} = C' \frac{Z^4 \lambda^3}{A}, \qquad (48.13)$$

où $C' = CN_{Av}$ est une nouvelle constante dont la valeur varie par saut.

Sur le schéma de la figure 90, on a représenté, pour l'argent et le cuivre, la variation du coefficient d'absorption massique, en unités arbitraires, en fonction de la longueur d'onde λ. En ordonnées on a porté la quantité $\tau_m^{1/3}$. La dépendance est partout linéaire, mais pour $\lambda = 0,0485$ nm (argent) et pour $\lambda = 0,13785$ nm (cuivre) on observe les arêtes d'absorption marquant l'arrêt de l'excitation du niveau K. Si on prolongeait la courbe d'absorption de l'argent du côté des ondes longues, on y aurait discerné trois « dents de scie » pour les longueurs d'onde de 0,3245, 0,3506, 0,3993 nm, correspondant aux nouvelles arêtes d'absorption (le niveau L se compose de trois sous-niveaux, voir fig. 88).

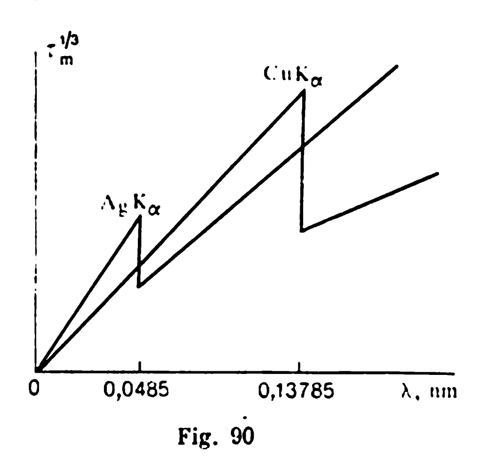


Fig. 90

L'absorption des rayons X est une *propriété* purement *atomique*. Le coefficient d'absorption moléculaire s'obtient par *addition* des coefficients d'absorption atomiques des éléments entrant dans la composition de la molécule. Connaissant les coefficients d'absorption atomiques, on peut donc calculer les coefficients d'absorption moléculaires de toute la multitude de composés chimiques. Selon la formule (48.12), les coefficients d'absorption atomiques augmentent rapidement avec le numéro atomique des éléments, proportionnellement à Z^4. Ces particularités de l'absorption des rayons X sont utilisées pour la radioscopie (cf. problème 1 à la fin du paragraphe).

10. Pour conclure, examinons encore un phénomène qui se produit dans l'enveloppe électronique de l'atome et qui détermine également une absorption des rayons X. Il est connu sous le nom d'*effet Auger* (né en 1899) ou *effet d'autoionisation des atomes excités* et a pour cause la redistribution de l'énergie dans l'atome excité. A la différence de la photoionisation usuelle, lorsque l'absorption d'un quantum X provoque l'expulsion d'un électron hors de l'atome, l'effet Auger se produit en deux étapes successives. Au cours de la pre-

mière étape, l'absorption d'un quantum X détermine l'excitation de l'atome consistant en l'expulsion d'un électron de la couche K où apparaît ainsi un trou. Lors de la seconde étape, ce trou est occupé par un électron provenant de la couche L. L'excès d'énergie $\mathcal{E}_K - \mathcal{E}_L$ peut être libéré sous forme d'un quantum de radiation caractéristique, comme on l'a vu plus haut, ou bien peut être utilisé pour expulser un électron des couches supérieures de l'atome. C'est cette deuxième éventualité qui constitue l'effet Auger; l'électron expulsé est dit *électron Auger*. Ainsi le trou qui a apparu dans la couche L subsiste et un second trou apparaît dans l'une des couches supérieures; cela signifie que l'atome se trouve ionisé deux fois.

C'est en photographiant, dans la chambre de Wilson, les photoélectrons extraits des gaz rares lourds (krypton par exemple) par les rayons X qu'Auger constata que dans certains cas *deux traces électroniques* avaient une origine commune. L'un des électrons était libéré par photoionisation de la couche K, tandis que l'autre était libéré par suite d'une redistribution interne de l'énergie de l'atome excité, i.e. par effet Auger.

Les électrons Auger emportent avec eux une part importante de l'énergie de l'atome excité. L'effet Auger se manifeste avec une grande intensité lorsque les transitions avec émission de photons sont interdites (par exemple transitions 0-0).

PROBLÈMES

1. Comparer les coefficients d'absorption des rayons X par les os et les tissus du corps humain. Le constituant principal des os est le phosphate de calcium $Ca_3(PO_4)_2$. L'absorption des rayons X par les tissus est déterminée surtout par l'eau H_2O.

Solution. En appliquant la formule (48.12) et compte tenu des valeurs des numéros d'ordre des éléments, on trouve

$$\frac{(\tau_a)_{os}}{(\tau_a)_{tissu}} = \frac{3\cdot 20^4 + 2\cdot 15^4 + 8\cdot 8^4}{2\cdot 1^4 + 8^4} \approx 150.$$

Selon (49.13), pour calculer les rapports des coefficients d'absorption massiques, on doit tenir compte du rapport des masses moléculaires des composés chimiques considérés. Pour $Ca_3(PO_4)_2$ on a $A = 3\cdot 40 + 2\cdot 30 + 8\cdot 16 = 308$. Pour H_2O $A = 2\cdot 1 + 16 = 18$. Par conséquent,

$$\frac{(\tau_m)_{os}}{(\tau_m)_{tissu}} \approx 150\,\frac{18}{308} \approx 9,$$

ce qui signifie que le coefficient d'absorption massique des os est près de 10 fois plus grand que celui des tissus. En raison de la différence des masses volumiques, le rapport des coefficients d'absorption $(\tau)_{os}$ et $(\tau)_{tissu}$ doit être multiplié par 2 ou 3. On comprend maintenant pourquoi sur les clichés des rayons X on discerne si nettement les os.

2. On sait que dans les atomes des terres rares, la couche interne f commence à être occupée après que les couches externes auront été remplies. Par conséquent, dans le nuage électronique de l'atome, on trouve des sites non occupés, i.e. des trous. Question : l'occupation de ces trous peut-elle donner lieu à l'émission spontanée de rayons X ?

R é p o n s e. Une émission spontanée de rayons X ne peut pas se produire car l'énergie de l'atome s'accroît (au lieu de diminuer) lorsqu'un électron appartenant à une couche externe vient se loger dans le trou.

§ 49. L'atome d'hélium

1. Après l'hydrogène les plus simples sont l'atome d'hélium ($Z = 2$) et les ions L^+ ($Z = 3$), Be^{++} ($Z = 4$), B^{+++} ($Z = 5$), . . . dont les enveloppes électroniques contiennent deux électrons chacune. Dans les spectres de ces atomes et de ces ions à deux électrons, on retrouve les mêmes séries spectrales que chez les atomes des métaux alcalins, mais chaque série se présente en deux exemplaires: on trouve deux séries principales, deux séries nettes, deux séries diffuses, etc. Dans l'un des exemplaires de chaque série toutes les raies sont simples (singulets) et dans l'autre toutes les raies sont des triplets, i.e. chaque raie se compose de trois raies rapprochées.

La raie la plus connue du spectre de l'hélium est la raie jaune D_3 qui fut observée pour la première fois en 1867 dans le spectre du Soleil et marqua la découverte de cet élément. C'est un triplet dont les longueurs d'onde sont : 587,5963 ; 587,5643 et 587,5601 nm, et dont les intensités sont dans le rapport 1 : 3 : 5. La distance de séparation des deux dernières raies n'est que de 0,0042 nm et c'est pour cela que pendant longtemps on n'arrivait pas à les séparer et au lieu d'un triplet D_3 on pensait avoir affaire à un doublet. Le triplet D_3 est le premier terme de la première série latérale des triplets. La série principale des triplets de l'hélium se situe dans l'infrarouge. Les séries correspondantes des singulets se situent surtout dans l'ultraviolet.

Entre les niveaux énergétiques singulets et triplets de l'hélium aucune transition quantique ne peut avoir lieu (*interdiction des intercombinaisons*). Cette circonstance servit de prétexte pour avancer une hypothèse selon laquelle l'hélium serait composé de deux éléments distincts: l'*orthohélium* produisant les raies triples et le *parahélium* dont les raies spectrales seraient des singulets. On verra par la suite que cette hypothèse s'avéra fausse et que l'interdiction des intercombinaisons n'est pas une règle absolument rigoureuse. La preuve en est qu'on trouve dans le spectre de l'hélium une raie, à vrai dire la seule, d'une longueur d'onde égale à 591,6 nm qui résulte de la transition du niveau triplet $^3P_{1/2}$ au niveau singulet $^1S_0$.

2. En mécanique quantique, la théorie des atomes à plusieurs électrons ne se heurte à aucune difficulté de principe, quoique les calculs pratiques soient très ardus et ne peuvent être effectués que sur ordinateur. Dans le cas le plus simple de l'atome d'hélium et des ions qui lui sont analogues, si on ne tient pas compte des spins, le problème se ramène à la résolution de l'équation de Schrödinger pour deux électrons. C'est ce problème que l'on considérera en omettant toutefois les calculs correspondants. Pour l'état stationnaire l'équa-

tion de Schrödinger est

$$\hat{H}\psi \equiv (\hat{H}_1 + \hat{H}_2 + \hat{U}_{12})\psi = \mathscr{E}\psi, \qquad (49.1)$$

où la fonction d'onde ψ dépend des coordonnées des deux électrons, $\mathscr{E}$ est l'énergie de l'état stationnaire, $\hat{H}_1$ et $\hat{H}_2$ sont les hamiltoniens du premier et du deuxième électron pour le cas où ces électrons n'interagissent pas entre eux:

$$\hat{H}_1 = -\frac{\hbar^2}{2\mu}\nabla_1^2 - \frac{Ze^2}{2r_1}, \qquad (49.2)$$

$$\hat{H}_2 = -\frac{\hbar^2}{2\mu}\nabla_2^2 - \frac{Ze^2}{2r_2} \qquad (49.3)$$

et $\hat{U}_{12}$ est la partie de l'hamiltonien total, qui tient compte des interactions entre les électrons:

$$\hat{U}_{12} = \frac{e^2}{r_{12}}. \qquad (49.4)$$

Ici ∇_1^2 et ∇_2^2 sont les opérateurs de Laplace pour le premier et le second électron:

$$\nabla_1^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}, \quad \nabla_2^2 = \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y_2^2} + \frac{\partial^2}{\partial z_2^2}.$$

On désigne par $\boldsymbol{r}_1$ (x_1, y_1, z_1) et $\boldsymbol{r}_2$ (x_2, y_2, z_2) les rayons vecteurs et les coordonnées cartésiennes du premier et du deuxième électron, et par r_{12} la distance entre les électrons. La charge du noyau est dénotée par Z (pour l'hélium $Z = 2$). Dans l'approximation considérée le noyau de l'atome est supposé fixe et on le prend pour origine des coordonnées.

3. Le problème qui vient d'être formulé est analogue au problème classique des trois corps de la mécanique céleste. Ce problème concerne le mouvement de deux planètes dans le champ gravitationnel du Soleil, compte tenu des interactions gravitationnelles entre les planètes. Quoique ce problème admette en principe une solution analytique sous forme de séries, ces séries sont absolument inutilisables pour les calculs pratiques (voir note en bas de la page 100). En mécanique céleste on a élaboré d'excellentes méthodes de calcul approché satisfaisant pleinement à toutes les exigences de l'astronomie pratique. Ces méthodes se fondent sur la *théorie des perturbations* qui met à profit le fait que l'interaction mutuelle des planètes est petite devant l'interaction de chacune des planètes avec le Soleil. Dans l'approximation d'ordre zéro on peut négliger l'interaction mutuelle des planètes. En utilisant la solution obtenue dans l'approximation d'ordre zéro, on peut ensuite tenir compte de l'interaction mutuelle des planètes en première approximation, puis trouver la deuxième approximation, etc. La mécanique quantique utilise la même théorie des perturbations pour résoudre le problème de l'atome à deux

électrons. Dans l'approximation d'ordre zéro on néglige aussi l'interaction mutuelle des électrons, quoique la situation soit ici notablement moins favorable qu'en mécanique céleste, puisque l'interaction mutuelle des électrons est loin d'être négligeable devant l'interaction de chacun des électrons avec le noyau atomique. Néanmoins, les résultats que l'on obtient ainsi sont presque satisfaisants et c'est ce qui justifie la mise en œuvre de la théorie des perturbations.

La principale difficulté du problème de l'atome d'hélium est liée à la présence, dans l'équation (49.1), du terme $\hat{U}_{12}\psi$ qui dépend des coordonnées des *deux* électrons. C'est pour cela que l'équation (49.1) ne peut avoir de solutions à variables séparables. En théorie des perturbations on considère le terme $\hat{U}_{12}\psi$ comme une « petite » correction qu'on rejette purement et simplement dans l'approximation d'ordre zéro. Ainsi dans l'approximation d'ordre zéro l'équation se présente sous la forme

$$\hat{H}^0\psi^0 = \mathcal{E}^0\psi^0, \tag{49.5}$$

où $\hat{H}^0$ désigne l'hamiltonien dans l'approximation d'ordre zéro, i.e. l'opérateur $\hat{H}^0 = \hat{H}_1 + \hat{H}_2$, où le terme perturbateur $\hat{U}_{12}$ a été rejeté. D'une façon analogue ψ^0 et $\mathcal{E}^0$ dénotent dans l'approximation d'ordre zéro la fonction d'onde et la valeur propre de l'énergie. Après avoir trouvé ψ^0 et $\mathcal{E}^0$ on cherche une solution en première approximation $\psi = \psi^0 + \psi^1$, $\mathcal{E} = \mathcal{E}^0 + \mathcal{E}^1$, en utilisant l'équation

$$(\hat{H}^0 + \hat{U}_{12})(\psi^0 + \psi^1) = (\mathcal{E}^0 + \mathcal{E}^1)(\psi^0 + \psi^1)$$

qui s'écrit, en vertu de (49.5),

$$\hat{H}^0\psi^1 + \hat{U}_{12}\psi^0 + \hat{U}_{12}\psi^1 = \mathcal{E}^1\psi^0 + \mathcal{E}^0\psi^1 + \mathcal{E}^1\psi^1.$$

On doit considérer les quantités $\hat{U}_{12}\psi^1$ et $\mathcal{E}^1\psi^1$ comme des « petits » d'ordre supérieur que l'on rejette en première approximation. On trouve en première approximation

$$(\hat{H}^0 - \mathcal{E}^0)\psi^1 = (\mathcal{E}^1 - \hat{U}_{12})\psi^0. \tag{49.6}$$

C'est une équation non homogène dont on connaît le second membre. On démontre aisément que l'équation (49.6) ne possède une solution que pour des valeurs du paramètre $\mathcal{E}^1$ pour lesquelles le second membre est orthogonal à la fonction d'onde ψ^0 de l'approximation d'ordre zéro, i.e.

$$\int \psi^{0*}(\mathcal{E}^1 - \hat{U}_{12})\psi^0\, d\tau = 0,$$

l'intégration étant étendue au volume à six dimensions $d\tau = dx_1\, dy_1\, dz_1\, dx_2\, dy_2\, dz_2$ des deux électrons. Si la fonction ψ^0 a été

normée à l'unité, on obtiendra

$$\mathcal{E}^1 = \int \psi^{0*} U_{12} \psi^0 \, d\tau, \tag{49.7}$$

ce qui signifie qu'en première approximation la correction $\mathcal{E}^1$ est égale à l'énergie potentielle d'interaction mutuelle des électrons U_{12} dont on calcule la moyenne à l'aide des fonctions d'onde d'approximation zéro. Après avoir trouvé $\mathcal{E}^1$ on peut résoudre l'équation (49.6) et trouver ψ_1. On ne le fera pas ici.

En procédant de façon semblable, on peut trouver en approximation d'ordre deux $\mathcal{E} = \mathcal{E}^0 + \mathcal{E}^1 + \mathcal{E}^2$, $\psi = \psi^0 + \psi^1 + \psi^2$ en considérant les corrections $\mathcal{E}^2$ et ψ^2 comme des quantités d'ordre de petitesse encore plus élevé, etc.

4. Examinons maintenant l'équation (49.5), i.e. la solution en approximation d'ordre zéro. En détaillant, on l'écrit sous la forme

$$(\hat{H}_1 + \hat{H}_2) \psi^0 = \mathcal{E}^0 \psi^0, \tag{49.8}$$

où l'opérateur $\hat{H}_1$ ne dépend que des coordonnées du premier électron et l'opérateur $\hat{H}_2$ que des coordonnées du second électron. De ce fait la solution ψ^0 sera une solution à variables séparables:

$$\psi^0 = \psi^0(1)\, \psi^0(2), \tag{49.9}$$

où (1) désigne l'ensemble des coordonnées du premier électron et (2) l'ensemble des coordonnées du second électron. En substituant (49.9) dans (49.8) et en divisant le résultat par $\psi^0 = \psi^0(1)\, \psi^0(2)$, on obtient

$$\frac{\hat{H}_1 \psi^0(1)}{\psi^0(1)} + \frac{\hat{H}_2 \psi^0(2)}{\psi^0(2)} = \mathcal{E}^0.$$

Le premier terme du premier membre ne dépend que des coordonnées 1, tandis que le second terme ne dépend que des coordonnées 2. Comme la somme de ces termes $\mathcal{E}^0$ est constante, chacun des termes pris séparément doit être constant. Cela signifie qu'on a les équations

$$\hat{H}_1 \psi^0(1) = \mathcal{E}_1^0 \psi^0(1), \quad \hat{H}_2 \psi^0(2) = \mathcal{E}_2^0 \psi^0(2), \tag{49.10}$$

où $\mathcal{E}_1^0$ et $\mathcal{E}_2^0$ sont les constantes vérifiant la condition

$$\mathcal{E}_1^0 + \mathcal{E}_2^0 = \mathcal{E}^0. \tag{49.11}$$

Les deux équations (49.10) sont en fait identiques et ne diffèrent que par les notations des coordonnées des premier et second électrons ainsi que par les valeurs numériques des constantes $\mathcal{E}_1^0$ et $\mathcal{E}_2^0$ (puisque les électrons peuvent se trouver dans des états différents; si les états sont les mêmes $\mathcal{E}_1^0 = \mathcal{E}_2^0$). Chacune des équations (49.10) décrit un état stationnaire de l'électron dans le champ du noyau en

supposant qu'on néglige les interactions mutuelles des électrons. Dans cette hypothèse

$$\hat{H}_1 = -\frac{\hbar^2}{2\mu}\nabla_1^2 - \frac{Ze^2}{r_1}, \quad \hat{H}_2 = -\frac{\hbar^2}{2\mu}\nabla_2^2 - \frac{Ze^2}{r_2}. \tag{49.12}$$

Ainsi dans l'approximation d'ordre zéro le problème initial se trouve ramené au calcul des fonctions propres et des valeurs propres d'un atome hydrogénoïde à un électron. La solution de ce dernier problème est bien connue.

5. Déterminons d'abord dans l'approximation d'ordre zéro l'énergie d'ionisation complète de l'atome d'hélium neutre, i.e. le travail d'éloignement à l'infini de ses deux électrons, l'atome se trouvant dans son état fondamental. Pour l'atome d'hydrogène à l'état fondamental le travail d'éloignement à l'infini de l'électron est $\mu e^4/2\hbar^2 \approx 13{,}539$ eV. Pour un atome analogue à l'atome d'hélium ionisé une fois, ce travail doit être Z^2 fois plus grand (voir formule (13.8)). Par conséquent, selon la formule (49.11) le travail d'éloignement à l'infini des deux électrons d'un atome de type hélium, dans l'approximation d'ordre zéro, sera égal à

$$\mathscr{E}^0_{\text{ionis}} = 2Z^2\frac{\mu e^4}{2\hbar^2}. \tag{49.13}$$

Ce résultat est évident puisque dans l'approximation d'ordre zéro on néglige l'interaction mutuelle des électrons. Pour l'atome d'hélium, la formule (49.13) conduit à la valeur

$$\mathscr{E}_{\text{ionis}} = 108{,}3 \text{ eV}.$$

Connaissant la fonction d'onde (49.9) dans l'approximation d'ordre zéro, l'application de la formule (49.7), où on pose $U_{12} = e^2/r_{12}$, permet de trouver la correction à apporter à l'énergie d'ionisation complète d'un atome de type hélium pour obtenir le résultat dans l'approximation d'ordre un. Les calculs fournissent le résultat suivant:

$$(\mathscr{E}^0 + \mathscr{E}^1)_{\text{ionis}} = \left(2Z^2 - \frac{5}{4}Z\right)\frac{\mu e^4}{2\hbar^2}. \tag{49.14}$$

Les valeurs numériques sont consignées dans le tableau 7 (voir p. 343). Comme on pouvait s'y attendre, dans l'approximation d'ordre zéro, on observe un désaccord flagrant entre théorie et expérience: l'erreur atteint près de 40 % pour He et près de 10 % pour C^{++++}. Mais déjà l'approximation d'ordre un fournit des résultats qui sont en bon accord avec les données expérimentales. Cela est d'autant plus étonnant que l'énergie d'interaction mutuelle des électrons n'est pas petite devant l'énergie d'interaction des électrons avec les noyaux atomiques. De nombreux savants se sont livrés à des calculs de l'énergie d'ionisation et de l'énergie d'excitation de l'atome d'hélium à des approximations d'ordres supérieurs et ont élaboré des

méthodes de calcul meilleures que la théorie des perturbations. Laissant de côté cette question, on indique dans la dernière colonne du tableau 7 les résultats obtenus par Hylleraas (1898-1965) qui n'a tenu compte ni des valeurs finies des masses des noyaux ni des corrections imposées par la théorie de la relativité. L'accord de ces résultats avec l'expérience est extraordinaire.

6. Cherchons maintenant une explication au fait que dans les spectres de l'hélium et des ions semblables à l'hélium toutes les séries spectrales sont doublées, comme on l'avait déjà signalé au début de ce paragraphe. Remarquons tout d'abord que toutes les séries spectrales résultent de l'excitation d'un seul électron et non de l'excitation simultanée des deux électrons. La probabilité de ce deuxième processus étant beaucoup plus petite que celle du premier, on ne l'envisagera pas. Considérons dans l'approximation d'ordre zéro un état de l'atome où l'un des électrons n'est pas excité et l'autre se trouve à l'état excité. Dans l'approximation d'ordre zéro l'état du premier électron est décrit par la fonction d'onde $\psi_1^0(1)$ et l'état du second par la fonction d'onde $\psi_k^0(2)$. L'indice du bas désigne l'ensemble des trois nombres quantiques (n, l, m_l) caractérisant l'état de l'électron. A l'indice inférieur 1 correspond l'état normal (non excité) ($n = 1$, $l = 0$, $m_l = 0$). Pour le moment on négligera le spin de l'électron.

Dans l'approximation d'ordre zéro la fonction d'onde des deux électrons est représentée par le produit $\psi_1^0(1)\,\psi_k^0(2)$. Or cette fonction ne satisfait pas au principe d'indiscernabilité des particules identiques (des électrons). Si on permute les électrons 1 et 2, on obtiendra la fonction $\psi_k^0(1)\,\psi_1^0(2)$ qui décrit le même état de l'atome: il se produit une dégénérescence d'échange (cf. § 46). Une superposition linéaire des deux fonctions fournit les mêmes états quantiques que l'état initial. Cependant, du fait de l'indiscernabilité des électrons, pour toute la multitude d'états concevables, ne seront réalisés que les états décrits par des fonctions d'onde symétriques et antisymétriques:

$$\left.\begin{aligned}\psi_s^0(1,\ 2) &= \psi_1^0(1)\,\psi_k^0(2) + \psi_k^0(1)\,\psi_1^0(2),\\ \psi_a(1,\ 2) &= \psi_1^0(1)\,\psi_k^0(2) - \psi_k^0(1)\,\psi_1^0(2).\end{aligned}\right\} \qquad (49.15)$$

Le caractère symétrique ou antisymétrique des fonctions (49.15) se rapporte à l'opération de permutation des électrons, i.e. de la permutation des arguments de la fonction d'onde, la « position » n'étant définie que par les coordonnées spatiales, les coordonnées de spin n'y participant pas. Or le principe d'antisymétrie des fonctions d'onde que doivent vérifier les électrons implique les fonctions d'onde *complètes*, dépendant aussi bien des coordonnées spatiales que des coordonnées de spin des électrons. La « position » de l'électron dans la fonction d'onde est définie non pas par trois, mais par *quatre* coordonnées. Les fonctions d'onde spatiales, donc incomplètes,

(49.15), vu qu'elles satisfont au principe d'indiscernabilité de particules identiques, peuvent être utilisées pour la description quantique de l'état des atomes. Cette assertion est contenue dans la formulation initiale et simplifiée du principe de Pauli. Ce principe exclut la possibilité d'existence d'états d'un atome pour lesquels les deux électrons auraient les quatre mêmes nombres quantiques n, l, m_l, m_s, mais admet que les deux électrons aient trois nombres quantiques identiques, à condition que le quatrième ait des valeurs différentes. En particulier, il peut exister des états avec $k = 1$, lorsque les deux électrons se trouvent à l'état fondamental caractérisé par les valeurs suivantes des nombres quantiques $n = 1$, $l = 0$, $m_l = 0$. Pour décrire un tel état, on doit évidemment choisir parmi les deux fonctions (49.15) possibles, la fonction symétrique, puisque la fonction antisymétrique est alors identiquement nulle. Comme dans ce cas la dégénérescence d'échange ne se manifeste pas, à l'état fondamental ne correspondent qu'une seule fonction d'onde et un seul niveau d'énergie.

On peut adopter chacune des expressions (49.15) pour approximation nulle et trouver l'approximation d'ordre un par la méthode de la théorie des perturbations. En procédant ainsi, on trouve deux fonctions d'onde, l'une symétrique et l'autre antisymétrique par rapport aux coordonnées. On les dénotera par ψ_s (1, 2) et ψ_a (1, 2) en omettant l'indice supérieur 0. L'état décrit par la fonction d'onde antisymétrique ψ_a (1, 2) est appelé *état ortho* et l'état décrit par la fonction d'onde symétrique ψ_s (1, 2) est l'*état para*. Il n'existe donc pas deux sortes d'hélium, l'hélium ortho et l'hélium para étant un *seul et même élément chimique se trouvant dans des états quantiques différents*. Ainsi se trouve expliquée l'origine des deux systèmes de termes et des transitions spectrales correspondantes entre ces termes. On conçoit que si ces transitions s'accompagnent d'une émission de lumière, elles doivent satisfaire aux règles de sélection dont il a été question au § 39.

7. Connaissant les fonctions d'onde en approximation d'ordre zéro, on peut calculer à l'aide de la formule (49.7) la correction à apporter à l'énergie de l'atome en première approximation. Comme cette formule implique que la fonction d'onde est normalisée à l'unité, on doit commencer par normaliser à l'unité les deux fonctions (49.15) — la fonction symétrique et la fonction antisymétrique. Notons C_s et C_a les facteurs de normalisation constants. Pour faire un vrai calcul, il faut évidemment commencer par calculer ces facteurs ; nous ne le ferons pas puisqu'il ne s'agit ici que de mettre en évidence l'*aspect physique de la question*, et pour cela il n'est nul besoin de connaître les valeurs numériques de C_s et C_a. Il importe cependant de tenir compte de la différence entre C_s et C_a (s'ils étaient identiques, on pourrait omettre les facteurs de normalisation en admettant que les expressions (49.15) sont déjà normalisées).

Pour les états décrits par la fonction d'onde symétrique la correction à l'énergie est égale à

$$\mathcal{E}_s^1 = |C_s|^2 \int \psi_s^{0*} \frac{e^2}{r_{12}} \psi_s^0 \, d\tau = |C_s|^2 \int \frac{e^2}{r_{12}} |\psi_s^0|^2 \, d\tau_1 \, d\tau_2, \quad (49.16)$$

où $d\tau_1 = dx_1\, dy_1\, dz_1$, $d\tau_2 = dx_2\, dy_2\, dz_2$ et pour les états décrits par une fonction d'onde antisymétrique la correction à l'énergie est:

$$\mathcal{E}_a^1 = |C_a|^2 \int \psi_a^{0*} \frac{e^2}{r_{12}} \psi_a^0 \, d\tau = |C_a|^2 \int \frac{e^2}{r_{12}} |\psi_a^0|^2 \, d\tau_1 \, d\tau_2. \quad (49.17)$$

En y substituant les valeurs des fonctions d'ondes tirées de (49.15) et ayant effectué la multiplication en tenant compte de l'indiscernabilité des électrons, on trouve

$$\begin{aligned} \mathcal{E}_s^1 &= |C_s|^2 (I_c + I_{\text{éch}}), \\ \mathcal{E}_a^1 &= |C_a|^2 (I_c - I_{\text{éch}}), \end{aligned} \quad (49.18)$$

où on a utilisé les notations

$$\begin{aligned} I_c &= 2 \int \frac{e^2}{r_{12}} |\psi_1^0(1)|^2 \, |\psi_k^0(2)|^2 \, d\tau, \\ I_{\text{éch}} &= 2 \int \frac{e^2}{r_{12}} \psi_1^0(1)\, \psi_k^{0*}(1)\, \psi_k^0(2)\, \psi_1^{0*}(2)\, d\tau. \end{aligned} \quad (49.19)$$

Les formules (49.16) et (49.17) montrent que les énergies $\mathcal{E}_s^1$ et $\mathcal{E}_a^1$ sont essentiellement positives. La quantité I_c est aussi essentiellement positive, comme en témoigne la première formule (49.19). Par conséquent,

$$I_c + I_{\text{éch}} > 0, \quad I_c - I_{\text{éch}} > 0.$$

La quantité $I_{\text{éch}}$ est également positive comme le montrent les calculs réels où on utilise les expressions explicites des fonctions d'onde d'atomes hydrogénoïdes. Il s'ensuit de (49.18) que *les niveaux énergétiques de l'état para se trouvent au-dessous des niveaux énergétiques de l'état ortho*. C'est pour cela que *l'état normal de l'atome d'hélium est l'état para.*

Le terme I_c se laisse interpréter de façon concrète par la théorie classique. La charge de l'un des électrons est pour ainsi dire « étalée » dans l'espace avec une densité volumique $\rho_1 = e\, |\psi_1^0(1)|^2$ et la charge de l'autre est « étalée » avec une densité volumique $\rho_2 = e\, |\psi_k^0(2)|^2$. L'expression intégrée dans I_c se laisse interpréter comme l'énergie potentielle de répulsion des charges $\rho_1\, d\tau_1$ et $\rho_2\, d\tau_2$, et l'intégrale elle-même représente l'énergie potentielle d'interaction des deux charges étalées dans l'espace. Le terme $I_{\text{éch}}$ n'a pas d'interprétation classique, car c'est un terme essentiellement quantique. Pour faire image, on pourrait dire que $I_{\text{éch}}$ apparaît parce que chacun des électrons se trouve pour ainsi dire simultanément dans les états ψ_1^0 et ψ_k^0. La quantité $I_{\text{éch}}$ est appelée *énergie d'échange* et l'interaction

correspondante porte le nom d'*interaction d'échange*. C'est l'existence de l'énergie d'échange qui détermine l'abaissement des niveaux énergétiques de l'état para et le relèvement des niveaux d'énergie de l'état ortho.

Il ne faut surtout pas s'imaginer que l'énergie d'échange est une énergie spéciale. Elle n'apparaît que par suite de l'*indiscernabilité de principe* de deux états quantiques dérivant l'un de l'autre par permutation de deux particules identiques en interaction mutuelle. C'est pour cette raison que l'interaction d'échange est spécifique non seulement dans le cas de forces coulombiennes, mais se manifeste dans tous les systèmes de particules identiques quelconques, quelle que soit la nature des forces d'interaction s'exerçant entre ces particules. Par exemple, l'interaction d'échange existe entre les nucléons des noyaux atomiques et se manifeste dans la *saturation des forces nucléaires*. Les interactions d'échange déterminent le phénomène de ferromagnétisme.

8. Passons maintenant à l'explication de la division des termes spectraux de l'atome d'hélium et de celle des raies spectrales correspondantes en singulets et en triplets. Pour cela il faut tenir compte des spins électroniques. Dans l'approximation d'ordre zéro les composantes spatiale et de spin de la fonction d'onde totale se laissent séparer, de sorte qu'on peut la présenter sous forme d'un produit des fonctions d'onde spatiale et de spin. Evidemment, les fonctions d'onde de spin doivent satisfaire, exactement comme les fonctions d'onde spatiales, aux conditions de symétrie ou d'antisymétrie. Il ne peut donc y avoir que quatre fonctions d'onde de spin, à savoir

$$\begin{aligned} \varphi_s^{(1)} &= \varphi_+(1)\,\varphi_+(2)\ \uparrow\uparrow, \\ \varphi_s^{(2)} &= \varphi_-(1)\,\varphi_-(2)\ \downarrow\downarrow, \\ \varphi_s^{(3)} &= \varphi_+(1)\,\varphi_-(2) + \varphi_+(2)\,\varphi_-(1)\ \uparrow\downarrow, \\ \varphi_a^{(4)} &= \varphi_+(1)\,\varphi_-(2) - \varphi_+(2)\,\varphi_-(1)\ \uparrow\downarrow. \end{aligned} \tag{49.20}$$

Pour les distinguer des fonctions d'onde spatiales, les fonctions d'onde de spin sont désignées par la lettre φ. Le signe + placé auprès de la fonction sert à indiquer que la projection m_s du spin de l'électron sur la direction déterminée par le moment orbital est égale à +1/2, tandis que le signe moins signifie que cette projection est égale à —1/2. Pour rendre l'écriture plus concrète on l'indique par des fléchettes orientées placées à droite. La première fonction symétrique $\varphi_s^{(1)}$ signifie que les projections m_s des spins des deux électrons sont égales à +1/2; la deuxième fonction $\varphi_s^{(2)}$ signifie que les deux projections sont égales à —1/2. Dans la troisième fonction symétrique $\varphi_s^{(3)}$ les projections des spins des deux électrons sont antiparallèles. Ainsi, dans les deux premiers cas le spin total des deux électrons est égal à 1, et ses projections sont respectivement

égales à $+1$ et -1. Or si le spin total est égal à 1, ses projections sur une direction choisie peuvent avoir trois valeurs: $+1$, -1, 0. La valeur 0 est réalisée par la troisième fonction d'onde symétrique $\varphi_s^{(3)}$.

En ce qui concerne la fonction de spin antisymétrique $\varphi_a^{(4)}$, la projection du spin qui lui correspond est égale à zéro, et comme c'est la seule projection existante, le spin total de cet état est égal à zéro.

On peut combiner les quatre fonctions de spin (49.20) avec les fonctions spatiales ψ_s et ψ_a afin d'obtenir la fonction d'onde totale. Mais comme il s'agit d'un système de deux électrons, la fonction d'onde totale du système doit être antisymétrique par rapport à l'opération de permutation des particules dans la fonction d'onde totale, caractérisée par les coordonnées spatiales et de spin. A cette condition ne peuvent satisfaire que les quatre combinaisons suivantes:

$$\varphi_s^{(1)}\psi_a, \quad \varphi_s^{(2)}\psi_a, \quad \varphi_s^{(3)}\psi_a \quad \text{et} \quad \varphi_a^{(4)}\psi_s.$$

Comme dans le cas qui nous occupe la normalisation ne joue aucun rôle, on peut s'en passer. Jusqu'ici il n'a été question que de combinaisons d'approximation d'ordre zéro. Par suite de la symétrie de l'énergie d'interaction Ze^2/r_{12}, dans toutes les approximations suivantes, les fonctions d'onde que l'on obtient présentent le même type de symétrie. Les trois premières fonctions (symétriques) forment un triplet de spin total égal à 1, dont les projections ont pour valeurs $+1$, -1, 0. A ce triplet correspond l'état ortho. Les niveaux énergétiques et les raies spectrales de ce triplet, satisfaisant aux règles de sélection, éclatent en trois sous-niveaux et en trois composantes par suite de l'interaction spin-orbite. Dans cet état les atomes d'hélium ont des moments magnétiques différents de zéro et en présence d'un champ magnétique présentent l'effet Zeeman.

La fonction d'onde $\varphi_a\psi_s$ décrit l'état para. La projection du spin total est alors nulle et comme c'est un état singulet, le spin total de l'état para est nul. Les atomes d'hélium se trouvant à l'état singulet et ayant un moment magnétique nul, ne donnent pas d'effet Zeeman.

En ce qui concerne l'interdiction des intercombinaisons dont il a été question au début de ce paragraphe, précisons que cette interdiction est liée à la conservation du spin des électrons lors de l'émission de rayonnements et admet donc des exceptions (cf. § 39, 5).

PROBLÈMES

1. Connaissant l'énergie de l'ionisation complète de l'atome d'hélium neutre à l'état fondamental (voir tableau 7), calculer d'abord l'énergie d'ionisation unique de cet atome, puis l'énergie d'ionisation de l'ion formé He^+.

Solution. Après que l'atome d'hélium neutre aura perdu un électron, il se transforme en un ion hydrogénoïde avec $Z = 2$. L'énergie d'ionisation de cet ion est égale à 13,539 Z^2 = 54,156 eV. Il s'ensuit que pour enlever le premier

Tableau 7

Atome ou ion	Energie d'ionisation complète, eV			
	Valeurs expérimentales	Valeurs calculées		
		Approximation d'ordre zéro	Approximation d'ordre un	Selon Hylleraas
He	78,98	108,3	74,46	78,98
Li^+	198,04	243,7	192,9	198,03
Be^{++}	371,51	433,2	365,5	371,49
B^{+++}	599,43	676,9	592,3	599,40
C^{++++}	881,83	974,8	873,3	881,82

électron il faut fournir un travail égal à 78,98 — 54,16 = 24,82 eV et pour enlever le second, un travail égal à 54,156 eV.

2. Estimer l'ordre de grandeur de l'énergie d'ionisation unique de l'atome de lithium à l'état fondamental.

Solution. Pour faciliter l'évaluation on supposera que les deux électrons K de l'atome de lithium font écran pour le champ du noyau, ce qui se traduit par une diminution du nombre de charge jusqu'à $Z = 3 - 2 = 1$. Avec cette hypothèse le problème se ramène au calcul du travail d'ionisation d'un atome hydrogénoïde lors de l'éloignement d'un électron de la couche L. Ce travail est égal à $\frac{13{,}539}{n^2} = \frac{13{,}539}{4} = 3{,}38$ eV. La valeur expérimentale est égale à 5,37 eV (voir tableau 7).

3. Connaissant l'énergie d'ionisation unique de l'atome de lithium pris à l'état fondamental (voir problème précédent) ainsi que l'énergie d'ionisation complète de l'ion Li^+ (voir tableau 7), calculer l'énergie d'ionisation de l'atome neutre de lithium à l'état fondamental lorsqu'on lui enlève deux électrons.

Réponse. 81,56 eV.

§ 50. Liaison chimique. Molécule d'hydrogène

1. Initialement les chimistes étaient conduits à postuler l'existence de « forces chimiques » spéciales afin d'expliquer la formation de molécules à partir d'atomes, ainsi que les réactions chimiques entre les atomes et les molécules. En réalité il n'existe pas de forces chimiques spéciales, car elles se ramènent aux forces d'interaction électriques (coulombiennes) ordinaires entre les particules chargées, i.e. entre les électrons et les noyaux constituant les atomes. Mais pour ramener les forces dites chimiques aux forces d'interaction électriques il faut avoir recours à la mécanique quantique.

On distingue deux types de liaisons chimiques : la *liaison ionique* (ou *hétéropolaire*) et la *liaison covalente* (ou *homéopolaire*). La liaison ionique apparaît lorsqu'on peut se représenter la molécule comme composée de deux ions : un ion positif et un ion négatif (par exemple, la molécule NaCl se compose des ions Na^+ et Cl^-). Lorsqu'on n'y arrive pas, on dit que la liaison est homéopolaire. La théorie de la liaison hétéropolaire (qui est semi-empirique) a été forgée, et avec

succès, avant l'avènement de la mécanique quantique. La liaison homéopolaire ne se laisse pas interpréter sur une base classique. La raison en est que les forces déterminant la liaison homéopolaire sont des *forces d'échange* qui, par principe, exigent une approche quantique. Pour bien le montrer, considérons la molécule d'hydrogène qui est la molécule à liaison homéopolaire la plus simple de toutes. C'est en 1927 que Heitler (1904-1981) et F. London (1900-1954) ont trouvé une solution au problème de la molécule d'hydrogène en appliquant la théorie des perturbations. Plus tard ce problème a été traité à l'aide de méthodes mathématiques plus élaborées, ce qui a permis d'atteindre à une précision notablement meilleure. Néanmoins, la méthode de Heitler et London est beaucoup plus simple. On n'envisagera ici que le principe de cette méthode, en omettant les calculs, puisqu'on ne cherche qu'à mettre en évidence la nature physique de la liaison homéopolaire sans chercher à présenter des calculs théoriques des valeurs numériques des grandeurs qui la caractérisent.

2. En mécanique quantique le problème de la molécule d'hydrogène se ramène à la résolution de l'équation de Schrödinger pour les états stationnaires

$$\hat{H}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathscr{E}\psi(\boldsymbol{r}_1, \boldsymbol{r}_2), \tag{50.1}$$

où $\boldsymbol{r}_1(x_1, y_1, z_1)$ et $\boldsymbol{r}_2(x_2, y_2, z_2)$ dénotent les coordonnées du premier et du deuxième électron. L'hamiltonien du système est de la forme

$$\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{V}, \tag{50.2}$$

où

$$\begin{aligned} \hat{H}_1 &= -\frac{\hbar^2}{2\mu}\nabla_1^2 - \frac{e^2}{r_{1A}}, \\ \hat{H}_2 &= -\frac{\hbar^2}{2\mu}\nabla_2^2 - \frac{e^2}{r_{2B}}, \\ \hat{V} \equiv V &= -\frac{e^2}{r_{1B}} - \frac{e^2}{r_{2A}} + \frac{e^2}{r_{12}} + \frac{e^2}{R}. \end{aligned} \tag{50.3}$$

L'hamiltonien $\hat{H}_1$ comporte, en plus de l'opérateur énergie cinétique de l'électron 1, un second terme représentant l'énergie potentielle de ce même électron dans le champ coulombien du noyau A. L'hamiltonien $\hat{H}_2$ est semblable à $\hat{H}_1$. Quant à l'hamiltonien $\hat{V}$, celui-ci caractérise quatre interactions : l'interaction de l'électron 1 avec le noyau B, l'interaction de l'électron 2 avec le noyau A, l'interaction mutuelle des électrons et l'interaction mutuelle des noyaux. Ici r_{1A}, r_{1B}, r_{2A}, r_{2B} dénotent respectivement les distances des électrons 1 et 2 jusqu'aux noyaux A et B, r_{12} est la distance entre les électrons et R la distance entre les noyaux A et B. On peut poser que les noyaux A et B étant lourds, sont fixes (cette approximation est dite

adiabatique). Dans cette approximation ψ n'est fonction que des coordonnées des seuls électrons. La distance R entre les noyaux qui est essentielle dans le cas de la liaison homéopolaire, ne figure dans l'expression de la fonction d'onde ψ qu'en tant que paramètre.

Il va de soi que la fonction d'onde totale dépend aussi bien des coordonnées de position des électrons que de leurs spins. La fonction d'onde totale doit naturellement être antisymétrique par rapport à la permutation des électrons dans un espace caractérisé par les coordonnées cartésiennes et de spin. On peut cependant ne considérer que les *fonctions d'onde spatiales* et pour tenir compte des spins il faut décomposer le problème en deux problèmes: 1) dans le premier problème *la fonction de spin est antisymétrique*, ce qui implique, conformément aux résultats du paragraphe précédent, que la fonction spatiale est symétrique; 2) dans l'autre problème *la fonction de spin est symétrique* et la fonction spatiale est antisymétrique. Dans le premier cas le spin total des deux électrons est égal à zéro, ce qui signifie qu'il ne peut y avoir qu'un seul état (*état singulet par rapport au spin*) caractérisé par une fonction spatiale symétrique. Dans le second cas le spin total des deux électrons est égal à 1 et ce spin peut prendre trois orientations spatiales différentes (*état triplet par rapport au spin*).

3. On conçoit que le problème ainsi formulé ne peut avoir de solution analytique, et qu'il faut faire appel à des méthodes approchées. Pour bien montrer comment on y parvient dans le cadre de la théorie des perturbations, supposons que les deux atomes d'hydrogène ont été éloignés l'un de l'autre à grande distance. On peut alors admettre que chacun des atomes est indépendant et peut être caractérisé par sa fonction d'onde individuelle. Comme il s'agit de la liaison chimique d'atomes non excités, on doit admettre que les deux atomes se trouvent dans l'état *fondamental*. Si les atomes se trouvent suffisamment loin l'un de l'autre, les fonctions d'onde des électrons 1 et 2 ne se recouvrent presque pas, ce qui permet de poser que l'électron 1 appartient à « son noyau » A et que le « noyau étranger » B n'affecte presque pas l'état de l'électron 1. De même on peut poser que l'électron 2 appartient à « son noyau » B. C'est pour cette raison qu'on a dégagé de l'hamiltonien complet $\hat{H}$ du système considéré les termes $\hat{H}_1$ et $\hat{H}_2$ qui sont les hamiltoniens des atomes individuels. On les considérera comme *parties principales* de l'hamiltonien et on en tiendra compte dans l'approximation d'ordre zéro. Le terme V représentant *l'hamiltonien d'interaction des atomes*, sera considéré comme une *petite perturbation* dont on devra tenir compte dans l'approximation d'ordre un et dans les approximations d'ordres supérieurs. Cela tient d'une part à ce que les atomes sont largement séparés et d'autre part à ce que dans l'expression (50.3) relative à V les deux

premiers termes sont partiellement compensés par les deux autres termes vu qu'ils sont de signes contraires.

Pour l'état singulet, pour lequel la fonction de spin est antisymétrique, on obtient dans l'approximation d'ordre zéro une fonction d'onde spatiale symétrique à variables séparables

$$\psi_s = C_s [\psi_A(1)\psi_B(2) + \psi_A(2)\psi_B(1)], \tag{50.4}$$

où C_s est la constante de normalisation. ψ_A désigne la fonction d'onde de l'électron appartenant au noyau A et ψ_B la fonction d'onde de l'électron appartenant au noyau B. On admet que les fonctions ψ_A et ψ_B sont discernables quoique leurs expressions analytiques soient identiques. Cette discernabilité des fonctions ψ_A et ψ_B résulte de ce qu'elles dépendent de variables différentes: dans le premier cas les coordonnées sont repérées par rapport au noyau A et dans le second par rapport au noyau B. Dans l'état triplet, lorsque la fonction de spin est symétrique, on obtient une fonction d'onde spatiale antisymétriqu

$$\psi_a = C_a [\psi_A(1)\psi_B(2) - \psi_A(2)\psi_B(1)], \tag{50.5}$$

où C_a est une autre constante de normalisation.

On calcule aisément les constantes C_s et C_a, mais la seule chose qui nous importe est de savoir qu'elles sont différentes. Calculons, à titre d'exemple, la constante C_s. On admettra que les fonctions ψ_A et ψ_B sont déjà normalisées. Dire que la fonction (50.4) est normalisée signifie que $\int \psi_s^* \psi_s \, d\tau = 1$, soit

$$\begin{aligned}
|C_s|^2 \int [\psi_A^*(1)\psi_B^*(2) + \psi_A^*(2)\psi_B^*(1)] \times \\
\times [\psi_A(1)\psi_B(2) + \psi_A(2)\psi_B(1)]\, d\tau_1\, d\tau_2 = \\
= |C_s|^2 \int \psi_A^*(1)\psi_A(1)\, d\tau_1 \int \psi_B^*(2)\psi_B(2)\, d\tau_2 + \\
+ |C_s|^2 \int \psi_A^*(2)\psi_A(2)\, d\tau_2 \int \psi_B^*(1)\psi_B(1)\, d\tau_1 + \\
+ |C_s|^2 \int \psi_A^*(2)\psi_A(1)\psi_B^*(1)\psi_B(2)\, d\tau_1\, d\tau_2 + \\
+ |C_s|^2 \int \psi_A^*(1)\psi_A(2)\psi_B^*(2)\psi_B(1)\, d\tau_1\, d\tau_2 = 1,
\end{aligned}$$

où $d\tau = d\tau_1\, d\tau_2$, $d\tau_1 = dx_1\, dy_1\, dz_1$, $d\tau_2 = dx_2\, dy_2\, dz_2$. Les deux premières intégrales de la partie moyenne sont égales à l'unité du fait de la normalisation des fonctions ψ_A et ψ_B, tandis que les deux dernières intégrales sont égales par suite de l'indiscernabilité des électrons. Dénotons l'une de ces intégrales, la première par exemple, par S, autrement dit

$$S = \int \psi_A^*(2)\psi_A(1)\psi_B^*(1)\psi_B(2)\, d\tau_1\, d\tau_2. \tag{50.6}$$

On aura alors

$$|C_s|^2 (2 + 2S) = 1,$$

d'où

$$C_s = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{1+S}}, \tag{50.7}$$

à condition d'omettre un facteur de phase non essentiel dans l'expression de C_s. On trouve de même

$$C_a = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{1-S}}. \tag{50.8}$$

De toute évidence ce n'est qu'en approximation d'ordre zéro que les fonctions d'onde ψ_s et ψ_a sont représentées par les expressions (50.4) et (50.5) (l'indice 0 a été omis pour simplifier). Mais le caractère de symétrie des fonctions d'onde spatiales se conserve dans toutes les approximations.

4. La connaissance des fonctions d'onde d'approximation d'ordre zéro suffit pour trouver l'énergie en première approximation. Dans l'état singulet (état à fonction d'onde spatiale symétrique)

$$\mathcal{E}_s = \int \psi_s^* (\hat{H}_1 + \hat{H}_2 + V) \psi_s d\tau \tag{50.9}$$

et dans l'état triplet (état à fonction d'onde spatiale antisymétrique)

$$\mathcal{E}_a = \int \psi_a^* (\hat{H}_1 + \hat{H}_2 + V) \psi_a \, d\tau. \tag{50.10}$$

Puisque ψ_s, considérée comme fonction des coordonnées r_1, est la fonction propre de l'opérateur $\hat{H}_1$, on a $\hat{H}_1\psi_s = \mathcal{E}_1\psi_s$. De même $\hat{H}_2\psi_s = \mathcal{E}_2\psi_s$, où $\mathcal{E}_1$ et $\mathcal{E}_2$ sont les énergies des atomes isolés. Comme ces atomes sont identiques et se trouvent tous deux dans l'état fondamental, $\mathcal{E}_1 = \mathcal{E}_2 \equiv \mathcal{E}_0$, où $\mathcal{E}_0$ est l'énergie de l'atome à l'état fondamental. La même chose vaut pour la fonction ψ_a. Si on remarque encore que les fonctions ψ_s et ψ_a sont normalisées, on comprend que dans les deux expressions ci-dessus l'hamiltonien $(\hat{H}_1 + \hat{H}_2)$ détermine l'apparition du terme $2\mathcal{E}_0$. D'après (50.3) la quantité V contient le terme e^2/R qui dépend de la distance de séparation des noyaux mais ne dépend pas des coordonnées des électrons. Dans la théorie que l'on expose ce terme joue le rôle de paramètre et adjoint la quantité e^2/R aux énergies $\mathcal{E}_s$ et $\mathcal{E}_a$. Ainsi

$$\mathcal{E}_s = 2\mathcal{E}_0 + \Delta\mathcal{E}_s, \qquad \mathcal{E}_a = 2\mathcal{E}_0 + \Delta\mathcal{E}_a, \tag{50.11}$$

où

$$\Delta\mathcal{E}_s = \int \psi_s^* V \psi_s \, d\tau, \quad \Delta\mathcal{E}_a = \int \psi_a^* V \psi_a \, d\tau. \tag{50.12}$$

En substituant les expressions (50.4) et (50.5) pour ψ_s et ψ_a dans les formules (50.12) et en tenant compte de l'indiscernabilité des

électrons, on obtient

$$\Delta\mathscr{E}_s = \frac{K+S}{1+S}, \quad \Delta\mathscr{E}_a = \frac{K-S}{1-S}, \tag{50.13}$$

où on a utilisé la notation

$$K = \int V\psi_A^*(1)\,\psi_A(1)\,\psi_B^*(2)\,\psi_B(2)\,d\tau_1\,d\tau_2. \tag{50.14}$$

En première approximation les quantités $\Delta\mathscr{E}_s$ et $\Delta\mathscr{E}_a$ représentent les corrections à apporter à l'énergie $2\mathscr{E}_0$ des électrons par suite de l'interaction coulombienne des atomes. A l'énergie de cette

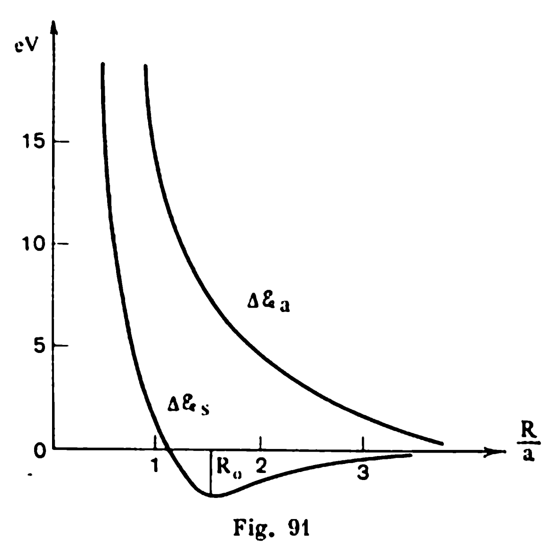


Fig. 91

interaction les noyaux contribuent par le terme e^2/R. L'interaction coulombienne entre les électrons présente un caractère dualiste, qui s'exprime d'une part par l'intégrale (50.14) et d'autre part par l'intégrale (50.6). Dans le premier cas l'électron 1 se trouve tout le temps dans le même état ψ_A et l'électron 2 dans le même état ψ_B. Leur interaction est alors de type classique, à condition de s'abstraire de l'« étalement » spatial des charges des électrons. Dans le second cas chacun des électrons se trouve simultanément dans l'état ψ_A et dans l'état ψ_B. Comme dans le cas de l'atome d'hélium, une telle interaction est un effet purement quantique, sans analogie en théorie classique. Cette interaction porte le nom d'*interaction d'échange* et a pour origine l'indiscernabilité des électrons. Dans les calculs l'indiscernabilité se traduit par l'utilisation des fonctions d'onde spatiales symétriques et antisymétriques.

5. Les intégrales K et S, de même que les énergies $\Delta\mathscr{E}_s$ et $\Delta\mathscr{E}_a$ sont des fonctions de la distance R entre les noyaux. La figure 91

représente la variation de ces énergies en fonction de la distance R. On porte en abscisses le rapport R/a, où a est le rayon de Bohr. Cette figure montre que lorsqu'on rapproche les atomes à l'état de spin singulet (spins antiparallèles) l'énergie d'interaction décroît d'abord puis augmente rapidement. Le minimum correspond à la valeur R_0 de la distance R. A l'état de spin triplet (spins parallèles) il n'y a pas de minimum d'énergie d'interaction.

Selon la figure 91, *la position d'équilibre des atomes ne se réalise que pour l'état de spin singulet*, lorsque $R = R_0$, distance de séparation pour laquelle l'énergie d'interaction $\Delta \mathscr{E}_n$ est minimale. *La formation d'une molécule n'est possible qu'à l'état singulet*, R_0 étant alors la distance d'équilibre entre les atomes dans la molécule. *A l'état triplet aucune molécule ne peut être formée.* Heitler et London ont trouvé en première approximation de la théorie des perturbations $R_0 = 1{,}51a \approx 0{,}080$ nm. La valeur expérimentale est $R_0 = 0{,}07395$ nm. On ne pourrait escompter un meilleur accord entre théorie et expérience, puisqu'en première approximation la théorie des perturbations n'est justifiée que pour $R \gg R_0$. Néanmoins même en première approximation elle donne une idée qualitative correcte de l'interaction des atomes d'hydrogène se trouvant dans les états singulet et triplet. Les méthodes de calcul variationnelles qui sont plus précises ont permis d'arriver, dans le cas d'une molécule d'hydrogène, à un accord quantitatif presque parfait entre les résultats théoriques et expérimentaux.

6. Des considérations qualitatives simples permettent d'expliquer la différence de comportement des atomes dans les états singulet et triplet. A l'état triplet la fonction d'onde spatiale (50.5) est antisymétrique et présente donc un nœud dans le plan de symétrie orthogonal à la droite reliant les noyaux A et B et passant par le milieu de cet intervalle, puisque dans ce plan par suite de l'indiscernabilité des électrons on a $\psi_A(1) = \psi_A(2)$, $\psi_B(1) = \psi_B(2)$. Pour cette même raison, à l'état singulet la fonction d'onde spatiale (50.4) devient maximale dans le même plan de symétrie. De ce fait dans l'état singulet, la densité de probabilité de présence des électrons à proximité du plan de symétrie est grande, tandis qu'elle est nulle dans l'état triplet. Par conséquent, à l'état singulet pour $R > R_0$ la répulsion coulombienne mutuelle des noyaux est compensée par l'action opposée des électrons, de sorte que, pour $R = R_0$, la résultante des forces d'interaction des atomes change de signe, ce qui signifie que la répulsion mutuelle cède la place à une attraction mutuelle. C'est la charge négative des électrons apparaissant entre les noyaux A et B qui les attache l'un à l'autre et produit ainsi un état lié. A l'état triplet, la probabilité de présence des électrons entre les noyaux étant petite, l'action compensatrice des électrons ne peut suffire à provoquer une attraction mutuelle des atomes.

Lorsque la distance entre les noyaux est petite, les atomes se

repoussent mutuellement, qu'ils se trouvent à l'état singulet ou à l'état triplet. Les forces de répulsion deviennent particulièrement grandes lorsque la distance de séparation des noyaux devient inférieure au rayon de Bohr a, puisque dans ce cas, au point médian entre les atomes, où les fonctions d'onde ψ_A et ψ_B se confondent, ces fonctions sont infiniment petites. (Selon la théorie semi-classique de Bohr, l'électron ne peut se rapprocher du noyau à une distance inférieure à a.) Aux petites distances, les forces de répulsion mutuelle des atomes sont déterminées moins par la répulsion coulombienne des protons que par la répulsion mutuelle des électrons. La répulsion coulombienne mutuelle des noyaux varie en raison inverse du carré de la distance entre les noyaux, tandis que selon les calculs théoriques, la répulsion coulombienne des électrons est régie par une loi de croissance exponentielle des forces de répulsion à mesure que diminue la distance entre les noyaux.

Les calculs montrent aussi que pour $R \gg a$, les forces d'attraction (pour l'état singulet) et de répulsion (pour l'état triplet) décroissent avec la distance selon une loi exponentielle.

Les formules (50.13) montrent que la différence quantitative entre les interactions mutuelles des atomes se trouvant à l'état singulet et à l'état triplet est déterminée par l'intégrale d'échange (50.6) qui dépend de R comme d'un paramètre. Les « forces chimiques » ne peuvent se manifester qu'à des distances R telles que l'expression intégrée dans l'intégrale d'échange (50.13) est notablement différente de zéro, donc dans la zone de « recouvrement » des fonctions d'onde électroniques $\psi_A(1)$ et $\psi_B(1)$ (ou bien des fonctions $\psi_A(2)$ et $\psi_B(2)$) des deux atomes.

7. Quoique la possibilité de formation d'une molécule neutre à partir d'atomes d'hydrogène à l'état singulet soit déterminée par le spin électronique, l'interaction de forces (magnétique) entre les spins des électrons n'y joue aucun rôle, car elle est très faible devant l'interaction d'échange. L'influence du spin ne se manifeste que dans la contrainte imposée à la fonction d'onde électronique complète du système de deux atomes d'hydrogène, celle-ci devant être antisymétrique par rapport à la permutation des électrons.

L'état lié de deux noyaux d'hydrogène (protons) peut apparaître même s'il n'y a qu'un seul électron et qu'il ne peut être question d'interaction d'échange. Cet état correspond à la molécule d'hydrogène ionisée H_2^+. La liaison est assurée par l'unique électron simultanément attaché aux deux noyaux. La possibilité d'apparition de forces d'attraction entre les noyaux peut être expliquée à l'aide d'un modèle classique. Supposons que l'électron négatif se trouve à mi-chemin entre les noyaux positifs (fig. 92). Les noyaux se repoussent avec une force $F_{++} = e^2/R$, tandis qu'entre l'électron et le noyau s'exerce une force d'attraction plus grande $F_{+-} = 4e^2/R$. Cet exemple classique ne peut qu'expliquer la cause d'apparition d'une

force d'attraction, mais la physique classique ne peut rien dire quant à la stabilité de l'état lié formé par les trois particules considérées. La mécanique quantique apporte une réponse positive à la question de la stabilité de l'état lié. Le problème se ramène à la résolution de l'équation de Schrödinger pour un électron se trouvant dans le champ électrique de deux noyaux (fixes) chargés positivement. On n'examinera pas ce problème.

8. Le cas où la liaison chimique des protons est réalisée par des muons négatifs, remplaçant les électrons, présente beaucoup d'inté-

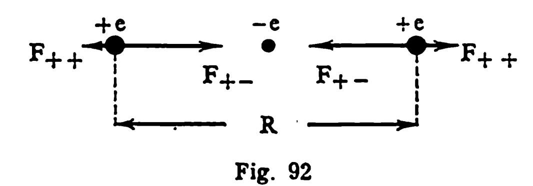


Fig. 92

rêt. On obtient alors une *molécule muonique* composée de deux protons et de deux muons négatifs. On peut remplacer l'un seulement ou les deux protons par un deuton ou un triton, vu que la charge de ces particules est la même que celle du proton. Le muon est une particule instable, dont la charge est égale à la charge de l'électron. Le muon se décompose en un électron et deux neutrinos, en moyenne au cours d'une durée de vie de $2 \cdot 10^{-6}$ s. C'est pendant cet intervalle de temps que peuvent exister les molécules muoniques. Comme la masse du muon est 207 fois plus grande que celle de l'électron, la distance moyenne entre les noyaux de la molécule muonique est 207 fois plus petite que dans la molécule d'hydrogène, c'est-à-dire $0{,}07395 : 207 = 3{,}57 \cdot 10^{-4}$ nm $= 3{,}57 \cdot 10^{-11}$ cm. Si les noyaux sont un proton et un deuton, à cette distance peut se produire une réaction nucléaire avec formation d'un noyau ^{3}He et une libération d'énergie $\sim 5{,}4$ MeV. Il s'ensuit que les muons peuvent catalyser cette réaction nucléaire. Ils peuvent assumer ce même rôle dans les réactions du deutérium avec du deutérium ou avec du tritium.

9. La construction d'une théorie quantitative satisfaisante de la liaison chimique homéopolaire dans les molécules complexes se heurte à de grandes difficultés mathématiques du fait du grand nombre d'électrons que comportent ces molécules. Une telle théorie n'existe pas encore. Pour expliquer les particularités qualitatives de la liaison chimique dans ces molécules on fait appel en chimie à des modèles fondés sur une extension de la théorie de la molécule d'hydrogène au cas de molécules complexes. Cette question ne sera pas envisagée dans ce cours.

§ 51. Hydrogène para et ortho

1. Les protons possèdent, comme les électrons, un spin 1/2 (en prenant $\hbar$ pour unité). Il s'ensuit que la fonction d'onde complète d'un système composé de ces particules doit être antisymétrique par rapport à la permutation de deux protons. Le moment magnétique de spin du proton est près de 657,5 fois plus petit que celui de l'électron. On peut donc affirmer, avec plus de raison encore que pour les électrons, que dans la molécule d'hydrogène la fonction d'onde complète des protons peut être représentée sous forme d'un produit des fonctions d'onde spatiale et de spin. Des raisonnements analogues à ceux utilisés pour l'étude d'un système de deux électrons montrent que dans la molécule d'hydrogène ne peuvent se réaliser que les deux états suivants:

1) L'état pour lequel la fonction d'onde de spin des protons est antisymétrique, ce qui implique que la fonction d'onde spatiale est symétrique. Dans ce cas l'hydrogène est appelé *parahydrogène*. Pour faire image, on dira que dans la molécule de parahydrogène les spins des protons sont antiparallèles.

2) L'état pour lequel la fonction d'onde de spin des protons est symétrique et la fonction spatiale est donc antisymétrique. On dira que dans ce cas les spins des protons sont parallèles et qu'on a affaire à l'*orthohydrogène*.

Comme dans la molécule d'hydrogène les spins électroniques se compensent toujours mutuellement, le spin I de la molécule n'est dû qu'aux spins des noyaux. Pour la molécule de parahydrogène $I = 0$, et pour la molécule d'orthohydrogène $I = 1$.

Pour définir la fonction d'onde spatiale des deux protons il est commode d'utiliser les coordonnées cartésiennes du centre de masse de la molécule, la distance de séparation des noyaux (des protons) et deux angles (coordonnées angulaires) définissant la direction de la droite reliant les noyaux. Comme les coordonnées du centre de masse ne changent pas lorsqu'on permute les protons, on peut ne pas en tenir compte. La distance entre les noyaux ne change pas non plus lorsqu'on les permute, mais elle varie par suite des vibrations mutuelles des noyaux le long de la droite qui les relie. L'énergie nécessaire pour exciter ces vibrations est assez importante (dans l'échelle des températures cette énergie équivaut à plusieurs milliers de kelvins). On supposera dans ce qui suit que les vibrations relatives des noyaux ne sont pas excitées (ne subsistent que les vibrations de zéro). Seules les coordonnées angulaires sont variables, ce qui signifie que les rotations de la molécule sont excitées. La permutation des protons équivaut à une rotation de la molécule de 180° autour de son centre de masse. Il s'ensuit que, dans le cas du parahydrogène, du fait d'une telle rotation la fonction d'onde spatiale des protons ne change pas, tandis que dans le cas de l'orthohydrogène, cette fonc-

tion change de signe. La forme explicite de la fonction d'onde angulaire des protons (qu'on s'abstient d'indiquer) laisse apparaître que dans le cas du parahydrogène cela se produit pour $L = 0, 2, 4, \ldots$, et dans le cas de l'orthohydrogène, pour $L = 1, 3, 5, \ldots$, L étant le nombre quantique caractérisant le carré du moment angulaire orbital de la molécule en rotation. Ainsi les niveaux rotatoires du parahydrogène correspondent à des niveaux quantiques *pairs* et ceux de l'orthohydrogène à des niveaux quantiques *impairs*. On en conclut que le niveau énergétique le plus bas d'une molécule de parahydrogène se trouve en dessous du niveau correspondant de la molécule d'orthohydrogène.

2. Pour transformer l'hydrogène para en ortho et vice versa, il faut modifier l'orientation des spins des noyaux. Or, par suite de la faible interaction des moments magnétiques de spin des noyaux, cette transformation est très peu probable lors des chocs des molécules d'hydrogène. Dans les conditions ordinaires, les transformations mutuelles des deux formes d'hydrogène ne se produisent pratiquement pas. Cela signifie que dans la molécule d'hydrogène les transitions quantiques sont régies par une règle de sélection qui, n'étant pas absolue, est cependant pratiquement rigoureuse

$$\Delta I = 0, \tag{51.1}$$

qui entraîne

$$\Delta L = \pm 2. \tag{51.2}$$

Cela signifie que *les transitions quantiques ne peuvent avoir lieu qu'entre des états quantiques rotatoires pairs ou impairs*. Si les teneurs du gaz en para et orthohydrogène sont différentes, cela se traduit par l'existence d'une alternance régulière des intensités des raies spectrales dans le spectre rotatoire (de bandes) de l'hydrogène moléculaire.

Ainsi, l'hydrogène moléculaire ordinaire se comporte comme un mélange de deux gaz ayant des propriétés légèrement différentes. Comme l'état rotatoire le plus bas ($L = 0$) correspond à un état para, à basse température la variété d'équilibre stable est le parahydrogène. A 20 K l'hydrogène à l'état d'équilibre est déjà du parahydrogène pratiquement pur. Mais un simple refroidissement de l'hydrogène jusqu'à cette température ne suffit pas pour y établir l'équilibre. Pour arriver à l'état d'équilibre, on utilise des catalyseurs (par exemple par désorption de l'hydrogène préalablement adsorbé par du charbon). Ce processus réalisé à 20 K permet d'obtenir du parahydrogène pratiquement pur. Ce dernier est très stable et peut être conservé dans des ballons en verre propres pendant plusieurs semaines.

A mesure que la température s'élève, le rapport des quantités de para et d'orthohydrogène se trouvant en équilibre varie. Comme il

n'existe pour le parahydrogène qu'un seul état quantique rotatoire ($I = 0$), tandis que l'orthohydrogène peut exister dans trois états (avec des spins —1, 0, +1), les poids statistiques de ces états sont dans le rapport de 1 : 3. A température élevée (pratiquement dès 300 K) le facteur exponentiel de Boltzmann cesse de jouer son rôle et les teneurs relatives d'hydrogène para et ortho dans le mélange d'équilibre sont dans le rapport de 1 : 3. L'hydrogène para se transforme en ortho en présence de catalyseurs (amiante platinée par exemple) à une température de quelques centaines de degrés Celsius ou encore sous l'action d'une décharge électrique. Mais comme il n'existe aucune température à laquelle l'état d'équilibre correspondrait à de l'orthohydrogène pur, celui-ci ne peut être obtenu. La plus petite teneur d'hydrogène para dans un mélange d'équilibre est égale à 25 %, conformément au rapport des poids statistiques des deux variétés d'hydrogène.

3. Lors d'un abaissement de température, la capacité calorifique de l'hydrogène moléculaire gazeux commence à diminuer dès la température ordinaire. A cette dernière température, la capacité calorifique molaire de l'hydrogène est égale à sa valeur classique $C_v = {}^5/_2R$ et tend vers la valeur classique $C_v = {}^3/_2R$ (cf. t. II, § 85) lorsque la température diminue. Cette diminution de la capacité calorifique est liée à une diminution de la contribution des mouvements rotatoires de la molécule. On n'arriva cependant à un accord complet entre théorie et expérience qu'après avoir tenu compte de ce que l'hydrogène est un mélange des variétés para et ortho. Les capacités calorifiques de ces constituants sont différentes parce que les niveaux rotatoires de l'orthohydrogène sont déplacés vers le haut par rapport aux niveaux rotatoires du parahydrogène. Le rapport d'équilibre entre para et orthohydrogène ne peut être atteint qu'au bout d'un temps long, qui est très grand par rapport au temps de chauffage et de refroidissement du gaz, qui est requis pour faire une mesure de sa capacité calorifique. A l'état hors d'équilibre la capacité calorifique de l'hydrogène gazeux dépendra des teneurs relatives en hydrogène ortho et para. Mais si on abandonne ce mélange hors d'équilibre à lui-même pour une durée suffisante, sa composition variera en fonction du temps, ce qui entraînera une variation lente dans le temps de sa capacité calorifique.

§ 52. Les forces moléculaires

1. Les forces d'interaction s'exerçant dans les molécules et les atomes se ramènent à des attractions et des répulsions coulombiennes entre les noyaux chargés positivement et les électrons chargés négativement. Parmi ces forces d'interaction citons en premier lieu les *forces chimiques* ou *forces de valence*, qui ont fait l'objet du § 50. Ces

forces qui s'exercent à petite distance sont liées à une orientation antiparallèle des spins des noyaux atomiques et dépendent du taux de recouvrement des fonctions d'onde électroniques des atomes interagissants. Comme la densité électronique décroît exponentiellement à mesure qu'augmente la distance jusqu'au centre de l'atome, les forces chimiques décroissent exponentiellement à mesure qu'augmente la distance de séparation des atomes. On supposera dans ce paragraphe que ni les atomes ni les molécules ne réagissent pas chimiquement entre eux, ce qui implique que les forces chimiques ne se manifestent pas.

D'autre part, on classe parmi les forces moléculaires les *forces de répulsion à petit rayon d'action*, qui apparaissent lorsque les enveloppes électroniques des particules interagissantes s'interpénètrent lors d'un rapprochement des particules. Ces forces décroissent exponentiellement très rapidement à mesure que la distance entre les particules augmente. C'est ce qui permet d'utiliser, en théorie cinétique des gaz, un modèle simplifié des molécules et des atomes en les assimilant à des billes élastiques solides dont les chocs mutuels font apparaître des forces de répulsion infiniment grandes.

Enfin, entre les atomes et les molécules neutres s'exercent encore des forces d'attraction. Ces forces décroissent en fonction de la distance non pas suivant une loi exponentielle, mais suivant une loi de puissance, donc beaucoup plus lentement. A la différence des forces à court rayon d'action, ce sont des *forces à grand rayon d'action*. Elles apparaissent par interaction des moments électriques dipolaires des atomes et des molécules et, d'une façon générale, chaque fois que leurs enveloppes électroniques sont déformées. Si les molécules sont polaires, i.e. possèdent des moments dipolaires permanents, on conçoit fort bien l'existence de ces interactions. Or les interactions existent et même surpassent en intensité les interactions des molécules polaires lorsque les molécules sont dénuées de moments dipolaires permanents. On dit dans ce cas qu'il s'agit des *forces de Van der Waals*, appelées aussi *forces de dispersion*.

Comment pourrait apparaître une interaction dipolaire entre des molécules dénuées de moment dipolaire? Des moments dipolaires variant dans le temps apparaissent chez les molécules et les atomes excités, mais si ces particules ne sont pas excitées, du point de vue classique, elles ne devraient pas avoir de moments dipolaires. Quant à la mécanique quantique, elle affirme, en accord avec l'expérience, que les atomes et les molécules non excités possèdent une énergie au point zéro à laquelle correspondent d'intenses vibrations de zéro. Les forces de Van der Waals tirent leur origine de ces vibrations. Les vibrations de zéro ne sont pas liées au mouvement thermique et ne dépendent donc pas de la température. Il s'ensuit que les forces de Van der Waals doivent être indépendantes de la température; l'expérience confirme cette conclusion.

2. Pour expliquer le rôle que jouent les vibrations de zéro dans l'apparition des forces de Van der Waals, ainsi que pour mettre en évidence leur nature quantique, il est indiqué de considérer non pas des atomes réels, mais leurs modèles idéalisés. Remplaçons chacun des atomes identiques interagissants par des oscillateurs harmoniques unidimensionnels de masse μ et de fréquence propre ω_0. Supposons que les oscillations se produisent suivant la droite reliant les oscillateurs. Lorsque ceux-ci s'écartent de x_1 et x_2 de leurs positions d'équilibre, ils acquièrent des énergies potentielles égales à $^1/_2\,\mu\omega_0^2x_1^2$ et $^1/_2\,\mu\omega_0^2x_2^2$, et des moments dipolaires $p_1 = ex_1$ et $p_2 = ex_2$. L'énergie potentielle d'interaction de ces dipôles est égale à $-p_1p_2/R^3 = -e^2x_1x_2/R^3$, R étant la distance entre les oscillateurs. En conséquence, l'hamiltonien du système des deux oscillateurs couplés est

$$\hat{H} = \hat{T}_1 + \hat{T}_2 + {}^1/_2\,\mu\omega_0^2(x_1^2 + x_2^2) - \lambda x_1x_2, \qquad (52.1)$$

où $\lambda = e^2/R^3$, et $\hat{T}_1$ et $\hat{T}_2$ sont les opérateurs énergie cinétique des oscillateurs, i.e.

$$\hat{T}_1 = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x_1^2}, \quad \hat{T}_2 = -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x_2^2}. \qquad (52.2)$$

Remplaçons x_1 et x_2 par les *coordonnées* dites *normales* q_1, q_2 d'un système d'oscillateurs couplés définies par les formules

$$x_1 = \frac{1}{\sqrt{2}}(q_1 + q_2), \quad x_2 = \frac{1}{\sqrt{2}}(q_1 - q_2). \qquad (52.3)$$

On écrira alors

$$\frac{\partial\psi}{\partial q_1} = \frac{\partial\psi}{\partial x_1}\frac{\partial x_1}{\partial q_1} + \frac{\partial\psi}{\partial x_2}\frac{\partial x_2}{\partial q_1} = \frac{1}{\sqrt{2}}\left(\frac{\partial\psi}{\partial x_1} + \frac{\partial\psi}{\partial x_2}\right).$$

On en déduit

$$\frac{\partial^2\psi}{\partial q_1^2} = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1^2} + 2\frac{\partial^2\psi}{\partial x_1\,\partial x_2} + \frac{\partial^2\psi}{\partial x_2^2}\right),$$

et

$$\frac{\partial^2\psi}{\partial q_2^2} = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial x_1^2} - 2\frac{\partial^2\psi}{\partial x_1\,\partial x_2} + \frac{\partial^2\psi}{\partial x_2^2}\right).$$

Il en découle

$$\frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}. \qquad (52.4)$$

On trouve pour l'énergie potentielle

$$\frac{1}{2}\mu\omega_0^2(x_1^2 + x_2^2) = \frac{1}{2}\mu\omega_1^2q_1^2 + \frac{1}{2}\mu\omega_2^2q_2^2,$$

où on a utilisé les notations

$$\omega_1^2 = \omega_0^2 - \lambda/\mu, \quad \omega_2^2 = \omega_0^2 + \lambda/\mu. \qquad (52.5)$$

Ainsi, l'hamiltonien d'un système d'oscillateurs couplés peut être mis sous la forme

$$\hat{H} = \hat{H}_1 + \hat{H}_2, \tag{52.6}$$

où

$$\begin{aligned} \hat{H}_1 &= -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial q_1^2} + \frac{1}{2}\mu\omega_1^2 q_1^2, \\ \hat{H}_2 &= -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial q_2^2} + \frac{1}{2}\mu\omega_2^2 q_2^2. \end{aligned} \tag{52.7}$$

Ce résultat signifie que dans les cas général les oscillations du système résultent de la superposition d'oscillations harmoniques de fréquences ω_1 et ω_2. La mécanique classique conduit évidemment au même résultat. Les oscillations de fréquences ω_1 et ω_2 sont appelées *oscillations normales* du système considéré. L'introduction des coordonnées normales permet de ramener le problème des oscillations d'un système couplé à celui des oscillations des deux oscillateurs *indépendants* ayant des fréquences ω_1 et ω_2 différentes.

3. L'énergie minimale du système, dite énergie au point zéro, correspond au cas où chacun des oscillateurs exécute des oscillations de zéro avec sa fréquence propre. L'énergie au point zéro du système est égale à

$$\mathscr{E}_0 = \frac{\hbar}{2}(\omega_1 + \omega_2). \tag{52.8}$$

Les fréquences ω_1 et ω_2 dépendent de la distance R entre les oscillateurs. Pour R grand

$$\begin{aligned} \omega_1 &= \omega_0\left(1 - \frac{\lambda}{\mu\omega_0^2}\right)^{1/2} \approx \omega_0\left(1 - \frac{\lambda}{2\mu\omega_0^2} - \frac{\lambda^2}{8\mu^2\omega_0^4}\right), \\ \omega_2 &= \omega_0\left(1 + \frac{\lambda}{\mu\omega_0^2}\right)^{1/2} \approx \omega_0\left(1 + \frac{\lambda}{2\mu\omega_0^2} - \frac{\lambda^2}{8\mu^2\omega_0^4}\right). \end{aligned}$$

Il en résulte que

$$\mathscr{E}_0 = \hbar\omega_0 - \frac{\hbar\lambda^2}{8\mu^2\omega_0^3} = \hbar\omega_0 - \frac{\hbar}{8\mu^2\omega_0^3}\frac{e^4}{R^6}, \tag{52.9}$$

ce qui signifie que l'énergie au point zéro d'oscillateurs couplés dépend de leur distance de séparation R et de ce fait elle joue le rôle de l'énergie potentielle d'interaction de ces oscillateurs. La quantité $\hbar\omega_0$ peut être considérée comme l'énergie au point zéro des oscillateurs non couplés. En rejetant cette constante additive non essentielle, on trouve pour l'énergie potentielle

$$U(R) = -\frac{\hbar e^4}{8\mu^2\omega_0^3}\frac{1}{R^6}. \tag{52.10}$$

Comme elle est négative, on en conclut que l'interaction mutuelle des oscillateurs présente un caractère d'attraction et ce, quelles que

soient les phases d'oscillation des oscillateurs. Les forces d'attraction correspondantes sont précisément les forces de Van der Waals s'exerçant entre les modèles d'atomes considérés. La nature quantique de ces forces se démontre par ce que la constante de Planck $\hbar$ figure dans la formule (52.10) *).

Dans le cas où les moments dipolaires des oscillateurs seraient mutuellement perpendiculaires, l'énergie potentielle de l'interaction dipolaire serait nulle et aucune force d'attraction ne pourra se manifester. Mais si ces moments forment entre eux un angle différent de 90°, on peut les décomposer en composants suivant la droite reliant les oscillateurs et une droite perpendiculaire à celle-ci, ce qui fait apparaître une attraction.

4. On peut transformer la formule (52.10) en y introduisant la polarisabilité β de l'atome dans un champ électrique continu à l'aide de la formule $p = ex = \beta E$. Comme dans le cas d'un oscillateur harmonique la force qui agit sur la charge e est égale à $\mu\omega_0^2 x$, à l'équilibre $\mu\omega_0^2 x = eE = ex/\beta$, d'où $e^2/(\mu\omega_0^2) = \beta$. Cette formule a été obtenue ici sur une base classique, mais elle peut être déduite de la formule de dispersion de la mécanique quantique (c'est d'ailleurs pour cette raison qu'on désigne les forces de Van der Waals sous le nom de forces de dispersion). On obtient finalement

$$U(R) = -\frac{\varepsilon\beta^2}{8}\frac{1}{R^6}, \tag{52.11}$$

où $\varepsilon = \hbar\omega$ est la différence entre les niveaux énergétiques adjacents de l'oscillateur harmonique.

Le résultat (52.11) ne change pas notablement lorsqu'on envisage des atomes réels. En deuxième approximation de la théorie des perturbations, la mécanique quantique conduit au résultat suivant:

$$U(R) = -k\frac{I\beta^2}{R^6}, \tag{52.12}$$

où I est l'énergie d'ionisation de l'atome et k un facteur numérique de l'ordre de l'unité.

---

*) L'attraction mutuelle des atomes idéalisés résulte de ce que l'interaction dipolaire des oscillateurs donne lieu à une diminution de leur énergie au point zéro totale. On peut le démontrer sans recourir à un développement en série de puissances des fréquences ω_1 et ω_2. En effet, on tire de (52.5)

$$\omega_1^2 + \omega_2^2 = 2\omega_0^2.$$

En assimilant les carrés des fréquences aux côtés d'un rectangle et d'un carré, on trouve, en vertu d'un théorème de géométrie bien connu, que $\omega_1^2\omega_2^2 < \omega_0^4$ ou $\omega_1\omega_2 < \omega_0^2$. Il s'ensuit que

$$\omega_1^2 + 2\omega_1\omega_2 + \omega_2^2 < 4\omega_0^2$$

ou

$$(\omega_1 + \omega_2)^2 < 4\omega_0^2, \quad \omega_1 + \omega_2 < 2\omega_0.$$

Or notre proposition se trouve démontrée par cette dernière inégalité.

En se plaçant à un point de vue différent, on peut expliquer la raison pour laquelle l'interaction dipolaire des molécules détermine leur attraction mutuelle. On observe une attraction lorsque les molécules se font face avec des charges de signes contraires (fig. 93) et une répulsion lorsqu'elles se font face avec des charges de même signe. Dans le premier cas, l'énergie potentielle d'interaction des

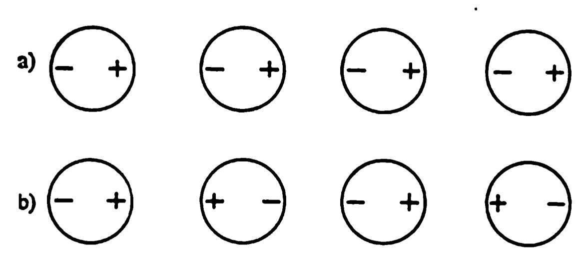


Fig. 93

molécules, qui est fonction de leur orientation réciproque, est minimale, et dans le second cas, elle est maximale. La première position correspond à un équilibre stable, la seconde à un équilibre instable. Selon la formule de Boltzmann, à l'état d'équilibre thermodynamique d'un gaz par exemple, la probabilité de la première position est notablement plus grande que celle de la seconde. Les molécules passent plus de temps dans la première position que dans la seconde et l'interaction résultante correspond à une attraction.

CHAPITRE VII

EXEMPLES D'EFFETS QUANTIQUES MACROSCOPIQUES

§ 53. Etats possibles d'une particule contenue dans un volume limité

1. Dans le t. II, § 82, où on avait exposé les fondements des statistiques quantiques de Bose-Einstein et de Fermi-Dirac pour le gaz parfait, on a obtenu des formules du nombre moyen de particules pouvant se trouver dans un seul état quantique à une température T donnée. Bien entendu, ne disposant encore d'aucune notion de mécanique quantique, on ne pouvait préciser ni la nature des états dans lesquels pourrait se trouver une particule, ni leur nombre dans un intervalle donné d'énergie. Maintenant on peut traiter ces questions pour une particule non relativiste quelconque par le procédé que nous avons mis en œuvre dans le cas du photon (cf. t. IV, § 117).

Soit une particule contenue dans une boîte qu'on posera, pour simplifier, être un cube d'arête L et à parois imperméables. Chaque paroi est une barrière de potentiel en ce sens qu'à l'intérieur de la boîte l'énergie potentielle de la particule est constante et on la pose égale à zéro, mais augmente indéfiniment jusqu'à l'infini à mesure que la particule se rapproche des parois ; au-delà de la paroi l'énergie reste infiniment grande. On doit introduire ces hypothèses pour exclure toute éventualité de sortie de la particule de la boîte (cf. §§ 24 et 28).

On supposera d'abord que la particule est dénuée de spin. Son état stationnaire à l'intérieur de la boîte est décrit par la fonction d'onde ψ vérifiant l'équation de Schrödinger

$$\nabla^2\psi + k^2\psi = 0, \tag{53.1}$$

où

$$k^2 = \frac{2\mu\mathscr{E}}{\hbar^2} = \vec{p}^{\,2}/\hbar^2. \tag{53.2}$$

Sur les parois de la boîte la fonction ψ doit s'annuler afin que la particule ne puisse s'en échapper.

On peut trouver une solution particulière de l'équation (53.1) par la méthode de séparation des variables en posant $\psi = \psi_x(x) \times \times \psi_y(y)\psi_z(z)$; on aura alors

$$\nabla^2\psi = \psi_x''(x)\,\psi_y(y)\,\psi_z(z) + \psi_x(x)\,\psi_y''(y)\,\psi_z(z) + \psi_x(x)\,\psi_y(y)\,\psi_z''(z).$$

En portant cette expression dans l'équation (53.1) et en la divisant par ψ, on obtient

$$\frac{\psi_x''(x)}{\psi_x(x)}+\frac{\psi_y''(y)}{\psi_y(y)}+\frac{\psi_z''(z)}{\psi_z(z)}=-k^2.$$

Cette équation doit être vérifiée pour toutes les valeurs de x, y, z. Le premier terme n'étant fonction que de x, il ne dépend pas des valeurs de y et z. Si on fixe les valeurs de y et z, les deux derniers termes du premier membre de l'équation seront des constantes, de même d'ailleurs que le premier terme $\psi_x''(x)/\psi_x(x)$. En appliquant ce même raisonnement aux deux autres termes, on arrive aux équations

$$\frac{\psi_x''(x)}{\psi_x(x)}=-k_x^2,\quad \frac{\psi_y''(y)}{\psi_y(y)}=-k_y^2,\quad \frac{\psi_z''(z)}{\psi_z(z)}=-k_z^2, \tag{53.3}$$

où k_x, k_y, k_z sont des constantes vérifiant la relation

$$k_x^2+k_y^2+k_z^2=k^2. \tag{53.4}$$

Toutes ces constantes doivent être positives, sinon les conditions aux limites ne seront pas satisfaites, ainsi qu'on le constatera dans ce qui suit. La solution générale de la première équation (53.3) est de la forme

$$\psi_x(x)=a\sin k_xx+b\cos k_xx.$$

La constante b doit être égale à zéro, puisqu'en vertu de la condition aux limites $\psi_x(0)=b=0$. La condition aux limites sur une paroi de la boîte $x=L$ impose que $\psi_x(L)=a\sin k_xL=0$, de sorte que $\sin k_xL=0$. On trouve des relations analogues pour k_y et k_z. Par conséquent,

$$k_x=\frac{\pi}{L}n_x,\quad k_y=\frac{\pi}{L}n_y,\quad k_z=\frac{\pi}{L}n_z, \tag{53.5}$$

où n_x, n_y, n_z sont des nombres entiers: $n_x, n_y, n_z = 1, 2, 3, \ldots$ (leurs valeurs négatives ne donnent pas de nouvelles solutions linéairement indépendantes, tandis que les valeurs $n_x=n_y=n_z=0$ conduisent aux solutions triviales $\psi_x=0$, $\psi_y=0$, $\psi_z=0$ et donc $\psi=0$).

On a obtenu ainsi la solution particulière de (53.1)

$$\psi=\sin k_xx\sin k_yy\sin k_zz, \tag{53.6}$$

qui s'annule sur les parois de la boîte. La fonction d'onde correspondante dépendant du temps est une onde stationnaire. La superposition d'ondes stationnaires d'amplitudes constantes représentera l'expression générale de la fonction d'onde à l'intérieur de la boîte. A chaque triade des nombres entiers n_x, n_y, n_z correspond une onde stationnaire, donc un état quantique stationnaire de la particule.

2. Pour trouver le nombre $dZ(k)$ d'états stationnaires contenus dans l'intervalle des nombres d'onde compris entre k et $k + dk$, construisons un réseau spatial cubique dont les mailles sont des cubes de côté π/L et de volume π^3/L^3. Dans ce cas le nombre $dZ(k)$ sera égal au nombre de nœuds contenus dans l'interstice de l'octant positif compris entre des sphères de rayons k et $k + dk$; ce nombre sera donc égal au rapport du volume de cet interstice au volume d'une maille:

$$dZ(k) = \frac{1}{8}\frac{4\pi k^2\,dk}{\pi^3/L^3} = \frac{Vk^2\,dk}{2\pi^2}, \tag{53.7}$$

où $V = L^3$ est le volume de la boîte.

Pour les électrons (et toutes les particules à spin 1/2) on doit multiplier par deux l'expression (53.7), puisqu'à chaque fonction d'onde spatiale correspondent deux états de spin à spins antiparallèles. Pour les photons on doit aussi multiplier (53.7) par deux pour tenir compte des deux polarisations orthogonales. Dans ces cas

$$dZ_{el}(k) = dZ_{ph}(k) = \frac{Vk^2\,dk}{\pi^2}. \tag{53.8}$$

Cette dernière formule ne peut être justifiée à l'aide de l'équation (53.1), car l'équation de Schrödinger ne s'applique pas aux photons, mais on l'avait déjà établie dans le t. IV, § 117, à l'aide des équations de Maxwell lors de l'étude du rayonnement thermique.

On peut passer des nombres d'onde aux impulsions par la formule $p = \hbar k$, d'où $dp = \hbar\,dk$. En utilisant ces variables, on écrira

$$dZ_{el}(p) = dZ_{ph}(p) = \frac{Vp^2\,dp}{\hbar^3\pi^2}. \tag{53.9}$$

On peut utiliser aussi l'énergie $\mathscr{E}$ de la particule en qualité de variable. Mais comme les relations entre l'énergie et l'impulsion sont différentes, on obtient dans ce cas des expressions différentes pour les électrons et les photons. Pour les électrons $\mathscr{E} = p^2/(2\mu)$,

$$dZ_{el} = \frac{V\sqrt{2\mu^3\mathscr{E}}\,d\mathscr{E}}{\pi^2\hbar^3}. \tag{53.10}$$

Pour les photons $p = \mathscr{E}/c$,

$$dZ_{ph} = \frac{V\mathscr{E}^2\,d\mathscr{E}}{\pi^2c^3\hbar^3}. \tag{53.11}$$

§ 54. Théorie de la capacité calorifique des corps solides de Debye

1. Il a été indiqué dans le t. II, §§ 69 et 85 que la mise en œuvre de la théorie quantique permit à Einstein d'expliquer dès 1906 la diminution de la capacité calorifique des solides à proximité du zéro absolu de température. Einstein considérait le corps solide comme un ensemble de N particules *indépendantes* (oscillateurs har-

moniques) vibrant autour de leurs positions d'équilibre avec la même fréquence ω. L'énergie moyenne revenant à chaque degré de liberté est dans ce cas donnée par la formule de Planck

$$\bar{\varepsilon} = \frac{\hbar\omega}{e^{\hbar\omega/kT} - 1}, \tag{54.1}$$

dans laquelle a été omis le terme $\hbar\omega/2$ représentant l'énergie au point zéro de l'oscillateur. On doit cependant tenir compte de ce terme chaque fois que l'amplitude des vibrations joue un rôle essentiel, comme par exemple dans l'étude de la variation de la diffusion des rayons X en fonction de la température. Dans le problème de la capacité calorifique, l'énergie au point zéro ne joue aucun rôle puisqu'elle ne dépend pas de la température. C'est pour cela que nous l'omettrons dans ce qui suit.

A haute température, la formule (54.1) se réduit à l'expression classique $\bar{\varepsilon} = kT$ et par suite le problème de la capacité calorifique se ramène à la loi de Dulong et Petit. Aux basses températures la formule d'Einstein prédit une diminution de la capacité calorifique avec la température et à l'approche du zéro absolu elle doit tendre vers zéro, conformément au théorème empirique de Nernst. Toutefois, l'accord entre théorie et expérience n'est que qualitatif. Ainsi, d'après la formule d'Einstein, la capacité calorifique des corps solides doit diminuer avec la température selon une loi *exponentielle*, tandis que l'expérience fournit une décroissance plus lente, selon une loi *de puissance*. On pourrait penser que ce désaccord tient non pas à un défaut de principe de la théorie, mais à ce que le modèle utilisé est trop grossier. La théorie d'Einstein considère que les oscillateurs sont indépendants, mais il serait plus conforme à la réalité de considérer que les oscillateurs sont *couplés*. Dans ce dernier cas on verra apparaître dans le corps non plus des vibrations d'une seule fréquence, mais tout un *spectre de fréquences* de vibration ω_i. Le nombre de ces fréquences est égal à $3N$, donc au nombre de degrés de liberté des N particules composant le corps (il est évident que certaines de ces fréquences seront les mêmes).

Si on considère le corps solide comme un système de N particules couplées exécutant des vibrations harmoniques normales, son énergie moyenne sera donnée par la formule

$$\bar{\mathscr{E}} = \sum_{i=1}^{3N} \bar{\varepsilon}_i = \sum_{i=1}^{3N} \frac{\hbar\omega_i}{e^{\hbar\omega_i/kT} - 1}$$

Le nombre de vibrations normales de fréquences inférieures à ω est assurément discret, mais étant très grand, on peut l'évaluer en approximation par une fonction continue $Z(\omega)$. Le nombre de vibrations normales comprises dans l'intervalle de fréquences de ω à $\omega + d\omega$ est également très grand, mais on peut l'évaluer en approximation

par la différentielle de la même fonction dZ (ω). Dans cette approximation on peut remplacer la formule précédente par la formule

$$\bar{\mathcal{E}} = \int_0^{\omega_{max}} \frac{\hbar\omega}{e^{\hbar\omega/kT}-1} dZ(\omega), \tag{54.2}$$

où ω_{max} désigne la fréquence maximale des vibrations normales. Cette fréquence maximale est définie par la relation

$$Z(\omega_{max}) = 3N, \tag{54.3}$$

puisque le nombre total des vibrations normales est égal au nombre de degrés de liberté $3N$. Ainsi, en théorie quantique, le problème de la détermination de l'énergie moyenne d'un corps solide se ramène à la détermination des fréquences propres des vibrations normales, tandis qu'en théorie classique l'énergie moyenne ne dépend que du nombre total de degrés de liberté.

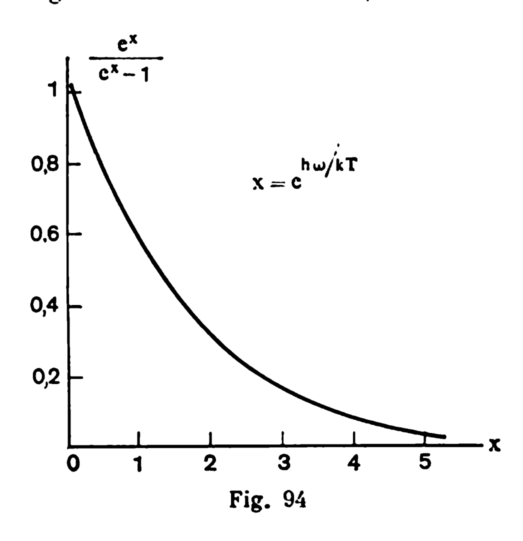


Fig. 94

2. Born et von Karman (1881-1963) se sont livrés au calcul des fréquences propres des vibrations du réseau cristallin dont la connaissance est requise par la théorie de la capacité calorifique. En général, c'est un problème des plus ardus, mais dans le cas de la capacité calorifique des corps solides aux basses températures, le problème peut être notablement simplifié en suivant la méthode utilisée par Debye. Cherchons à exprimer l'énergie moyenne $\bar{\varepsilon}$ de l'oscillateur, définie par la formule de Planck (54.1), en fonction de la température T. Pour cela, récrivons cette formule sous la forme

$$\bar{\varepsilon} = kT \frac{x}{e^x - 1}, \tag{54.4}$$

où $x = \hbar\omega/kT$ est une variable sans dimension. Le graphe de cette fonction représenté sur la figure 94 montre que dans l'expression (54.2) de l'énergie moyenne du corps ne comptent que les termes correspondant aux *basses fréquences* des vibrations normales. A ces basses fréquences correspondent des longueurs d'onde qui sont grandes devant le paramètre du réseau cristallin. Cela permet de négliger

la structure atomique du corps et d'assimiler les vibrations normales à des ondes stationnaires ultrasoniques ou infrasoniques dans un milieu élastique continu. Ce sont les mêmes ondes que celles qui déterminent la structure fine des raies spectrales résultant de la diffusion moléculaire de la lumière (effet Mandelstam-Brillouin, cf. t. IV, § 99). Il s'ensuit que les basses fréquences propres essentielles peuvent être calculées par les méthodes de la *théorie de l'élasticité* qui considère que le milieu est *continu*.

On arrive à établir l'expression de dZ (ω) à partir des équations différentielles de la théorie de l'élasticité exactement de la même façon que fut établie la formule (53.7) d'une grandeur semblable dans le cas des ondes de de Broglie. Il importe cependant de remarquer que dans un corps solide peuvent se propager des ondes sonores *longitudinales* et *transversales*. Dans une seule et même direction ne peut se propager qu'une seule onde sonore longitudinale d'une fréquence déterminée. Par conséquent, pour les ondes longitudinales, on peut conserver sans changement la formule (53.7) (en remplaçant toutefois la vitesse de phase des ondes de de Broglie par la vitesse du son). Mais il peut y avoir deux ondes transversales se propageant suivant une direction donnée avec une même fréquence. Pour ces ondes on doit multiplier l'expression (53.7) par deux. On obtient ainsi

$$dZ(\omega) = \frac{V\omega^2\,d\omega}{2\pi^2}\left(\frac{1}{c_{\parallel}^3} + \frac{2}{c_{\perp}^3}\right) = \frac{3V\omega^2\,d\omega}{2\pi^2 c^3} \tag{54.5}$$

où V est le volume du corps, $c_{\parallel}$ la vitesse de propagation des ondes sonores longitudinales, et $c_{\perp}$ la vitesse de propagation des ondes sonores transversales. La quantité c est une certaine « vitesse moyenne » définie par la relation

$$\frac{3}{c^3} = \frac{1}{c_{\parallel}^3} + \frac{2}{c_{\perp}^3} \tag{54.6}$$

Nous n'avons pas tenu compte de l'anisotropie des propriétés élastiques des cristaux qui existe même dans le cas de cristaux cubiques. On a supposé que le corps était isotrope et que ses propriétés élastiques pouvaient être caractérisées par deux constantes, par exemple les vitesses du son $c_{\parallel}$ et $c_{\perp}$. La prise en compte de l'anisotropie ne change pas grand-chose aux résultats et elle n'est pas justifiée dans le cadre de la méthode approchée de Debye.

3. Selon (54.2) l'énergie moyenne du cristal est

$$\bar{\mathscr{E}} = \frac{3V\hbar}{2\pi^2 c^3} \int_0^{\omega_{max}} \frac{\omega^3\,d\omega}{e^{\hbar\omega/kT} - 1}\,; \tag{54.7}$$

en y introduisant la notation déjà utilisée $x = \hbar\omega/kT$ et la notation $x_{max} = \hbar\omega_{max}/kT$, on obtient

$$\bar{\mathscr{E}} = \frac{3Vk^4T^4}{2\pi^2c^3\hbar^3} \int_0^{x_{max}} \frac{x^3dx}{e^x-1}. \tag{54.8}$$

A basse température l'expression sous le signe d'intégration est très petite pour les fréquences élevées ($x \gg 1$). Dans ce cas, la connaissance exacte de la limite supérieure x_{max} n'est pas nécessaire, et on peut la poser égale à l'infini, i.e.

$$\bar{\mathscr{E}} = \frac{3Vk^4T^4}{2\pi^2c^3\hbar^3} \int_0^{\infty} \frac{x^3\,dx}{e^x-1}. \tag{54.9}$$

L'intégrale figurant dans (54.9) est exactement la même que celle qui figurait dans la démonstration de la loi de Stefan-Boltzmann à partir de la formule de Planck (t. IV, § 118). Elle a pour valeur

$$\int_0^{\infty} \frac{x^3\,dx}{e^x-1} = \frac{\pi^4}{15}. \tag{54.10}$$

Par conséquent, à basse température ($x_{max} \gg 1$)

$$\bar{\mathscr{E}} = DT^4, \tag{54.11}$$

où

$$D = \frac{V\pi^2k^4}{10c^3\hbar^3}. \tag{54.12}$$

On trouve pour la capacité calorifique du corps

$$C_V = \left(\frac{\partial\bar{\mathscr{E}}}{\partial T}\right)_V = 4DT^3. \tag{54.13}$$

On voit donc qu'*à proximité du zéro absolu la capacité calorifique du réseau cristallin varie proportionnellement au cube de la température.* C'est la *loi des cubes* établie théoriquement par Debye. Selon cette loi, pour $T = 0$ la capacité calorifique devient égale à zéro, en accord avec le théorème de Nernst. La loi des cubes de Debye est en accord satisfaisant avec les résultats expérimentaux (voir cependant le sous-paragraphe 5). Il ne faut pas oublier que la formule (54.13) ainsi que la loi des cubes ne concernent que la capacité calorifique du réseau cristallin. Dans le cas des métaux, à la capacité calorifique du réseau vient s'ajouter la capacité calorifique des électrons libres, qui varie proportionnellement à la température (cf. t. II, § 85).

4. A haute température ($\hbar\omega/kT \ll 1$) la formule (54.7), compte tenu de (54.3), conduit au résultat classique $\bar{\mathscr{E}} = 3NkT$, qui est un résultat correct. Ce résultat ne découle évidemment pas de la for-

mule (54.5) qui cesse d'être valable aux fréquences élevées. Ce résultat est une conséquence de ce qu'en théorie classique la forme du spectre des fréquences des vibrations normales ne joue aucun rôle, puisque la seule chose qui compte est le nombre total de vibrations, celui-ci étant correctement déterminé par la formule (54.3).

Par conséquent, la formule (54.7) fournit des résultats corrects aux températures très élevées et aux très basses températures. On peut donc l'appliquer en tant que *formule d'interpolation* aux températures intermédiaires.

Déterminons maintenant la valeur de ω_{max} en nous fondant sur le modèle continu du corps. D'après la formule (54.3)

$$Z(\omega_{max}) = \int_0^{\omega_{max}} dZ(\omega) = \frac{3V}{2\pi^2 c^3} \int_0^{\omega_{max}} \omega^2 \, d\omega = \frac{V}{2\pi^2 c^3} \omega_{max}^3 = 3N,$$

d'où

$$\omega_{max} = \left(\frac{6\pi^2 c^3 N}{V} \right)^{1/3} = \frac{2\pi c}{a} \left(\frac{3}{4\pi} \right)^{1/3}, \qquad (54.14)$$

où a est le paramètre du réseau défini par $V = Na^3$. La longueur minimale de l'onde élastique correspondant à la fréquence maximale ω_{max} est

$$\lambda_{min} = \frac{2\pi c}{\omega_{max}} = a \left(\frac{4\pi}{3} \right)^{1/3} \approx 1{,}6a. \qquad (54.15)$$

Il va de soi que ce calcul ne fournit que l'ordre de grandeur des valeurs réelles de ω_{max} et λ_{min}, puisque pour des ondes aussi courtes le modèle continu des corps solides ne peut plus donner des résultats exacts.

Définissons la « température de Debye » par la formule

$$T_D = \frac{\hbar \omega_{max}}{k} = \frac{2\pi c \hbar}{ka} \left(\frac{3}{4\pi} \right)^{1/3}. \qquad (54.16)$$

On peut alors écrire la formule (54.8) pour une mole sous la forme suivante:

$$\bar{\mathscr{E}} = 9N_A kT \left(\frac{T}{T_D} \right)^3 \int_0^{T_D/T} \frac{x^3 \, dx}{e^x - 1} = 3RT \cdot 3 \left(\frac{T}{T_D} \right)^3 \int_0^{T_D/T} \frac{x^3 \, dx}{e^x - 1}, \quad (54.17)$$

où $R = Nk$ est la constante universelle des gaz. A haute température $T \gg T_D$ cette expression se réduit à la formule classique $\bar{\mathscr{E}} = 3RT$ qui exprime la loi de Dulong et Petit. Aux températures notablement plus petites que la température de Debye, on retrouve la formule (54.11) et la loi des cubes de Debye. En différentiant (54.17) par rapport à la température, on obtient la formule d'interpolation de De-

bye pour la capacité calorifique molaire des corps solides:

$$C_V = 3R\left\{12\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^3\,dx}{e^x-1} - \frac{3T_D/T}{e^{T_D/T}-1}\right\}. \qquad (54.18)$$

On indique dans le tableau 8 les valeurs des températures de Debye calculées à l'aide de la formule (54.16) et celles déterminées à

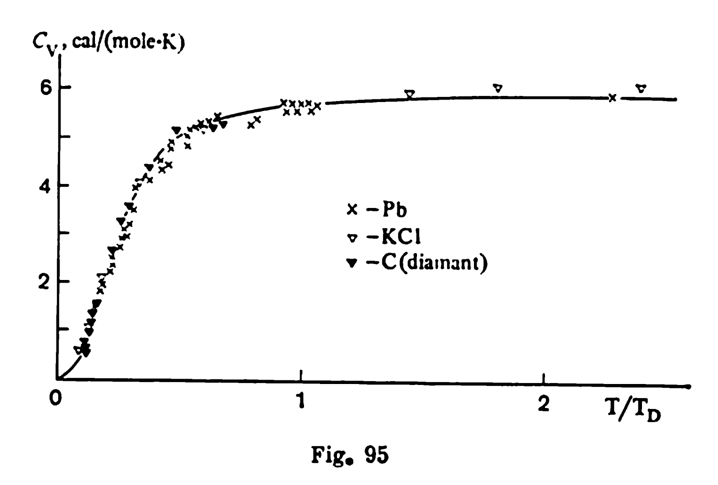


Fig. 95

partir de la courbe expérimentale de la capacité calorifique. La figure 95 représente la variation de la capacité calorifique en fonction du rapport T/T_D, calculée par la formule (54.18) (courbe en trait plein); les points qui y sont marqués correspondent aux valeurs expérimentales trouvées pour Pb, KCl et C (diamant).

Tableau 8

Valeurs des températures de Debye T_D

Substance	Valeurs expérimentales	Valeurs calculées d'après la formule (54.16)
Aluminium	398	402
Cuivre	315	332
Argent	215	214
Plomb	88	73
Diamant	1860	—

5. On doit remarquer que la théorie de Debye qui vient d'être exposée n'est valable que pour les cristaux composés de particules dont l'état interne est pratiquement indépendant de la température.

Cela implique que la structure interne de ces particules ne se manifeste pas dans la capacité calorifique. Dans la majorité des cas, à cette condition satisfont les cristaux composés d'atomes ou d'ions, pour lesquels la distance entre le niveau normal et le premier niveau excité est grande devant kT. Il existe cependant des atomes et des ions dont les niveaux énergétiques inférieurs sont très rapprochés. Par exemple, pour le sulfate de gadolinium cristallin, le niveau énergétique inférieur de l'ion gadolinium se compose de huit sous-niveaux dont les distances de séparation mesurées dans l'échelle de température correspondent à une température de Debye égale à 1,6 K. A très basse température $T \sim 7$ K on voit apparaître une capacité calorifique supplémentaire due à l'excitation de ces sous-niveaux. A ces basses températures la capacité calorifique du réseau est petite devant cette capacité calorifique supplémentaire. A la température $T = 1{,}6$ K la capacité calorifique du cristal est près de 500 fois plus grande que la capacité calorifique du réseau cristallin. Aux températures encore plus basses la capacité calorifique du cristal tombe naturellement à zéro.

Pour les cristaux composés de molécules complexes, on observe parfois une capacité calorifique liée au mouvement thermique des atomes ou des groupements atomiques au sein de la molécule. On peut poser en première approximation que les vibrations des molécules à l'intérieur du réseau cristallin n'affectent pas l'état interne des molécules. Dans ces cas la capacité calorifique du cristal se présente sous la forme

$$C = C_{\text{rés}} + C_{\text{int}}.$$

Dans certains cas, la contribution apportée à la capacité calorifique par le mouvement interne peut être notable. Par exemple, la capacité calorifique liée aux vibrations internes dans les molécules de benzène représente à $T \approx 150$ K près de 20 % de la capacité calorifique du réseau, et à $T \approx 270$ K elle atteint 80 %.

§ 55. Types de liaisons des atomes dans les corps solides

1. Dans le paragraphe précédent, on s'est abstrait de la structure atomistique interne des corps solides pour faire la théorie de leur capacité calorifique. Dans l'étude d'autres propriétés des solides, on ne peut procéder ainsi.

Les atomes, les molécules ou les ions qui composent les solides cristallins effectuent de petites vibrations autour des nœuds du réseau cristallin correspondant. Les propriétés géométriques des réseaux cristallins, i.e. leur symétrie, sont du ressort de la cristallographie. On en a donné un aperçu dans le ch. XII du tome II, aussi ne s'y arrêtera-t-on plus.

La disposition régulière des unités structurales (atomes, ions, molécules) des cristaux dans les nœuds du réseau cristallin demande à être expliquée au point de vue de l'atomistique. Il vient aussitôt à l'esprit que pour donner une explication adéquate, il faudrait résoudre l'équation de Schrödinger

$$\hat{H}\psi = \mathscr{E}\psi, \tag{55.1}$$

où la fonction d'onde ψ dépend des coordonnées de tous les noyaux et de tous les électrons du corps. Il va de soi que cette équation ne peut avoir, comme toujours, de solutions admissibles que pour certaines valeurs de l'énergie $\mathscr{E}$. La résolution de l'équation de Schrödinger permettrait de trouver les valeurs de $\mathscr{E}$ pour lesquelles le corps serait à l'état cristallin, ainsi que de déterminer la structure du réseau cristallin. Les électrons qui occupent les interstices entre les noyaux atomiques sont là pour maintenir les unités structurales dans des positions d'équilibre, i.e. dans les nœuds du réseau cristallin. La solution de l'équation devrait fournir toutes les structures possibles des réseaux cristallins. Elle doit embrasser toutes les transitions de phase entre eux, ainsi que les passages à l'état liquide et ensuite à l'état gazeux de la substance.

Or la résolution exacte de l'équation de Schrödinger d'un corps macroscopique renfermant un nombre énorme de variables est une tâche absolument irréalisable. C'est un problème encore plus compliqué que le problème de plusieurs corps de la mécanique classique. D'autre part, la forme explicite de la fonction ψ d'un corps macroscopique ne présente par elle-même aucun intérêt, mais il est possible et utile d'étudier la fonction ψ sans chercher à l'expliciter, en se fondant sur l'équation de Schrödinger, afin d'en déterminer les propriétés susceptibles de caractériser les particularités observées des corps macroscopiques, notamment l'existence même et la structure des réseaux cristallins. Cependant ce problème n'a jamais été posé, puisqu'on n'a aucune certitude que l'on pourrait obtenir ainsi les résultats escomptés. Dans les études concernant la structure d'un réseau cristallin et les forces qui retiennent les atomes, les ions ou les molécules dans leurs positions d'équilibre, on est obligé de s'appuyer largement sur les données expérimentales.

2. On distingue les *cristaux ioniques*, les *cristaux covalents*, les *cristaux moléculaires*, les *métaux*.

Le réseau cristallin des cristaux ioniques se compose d'ions positifs et négatifs entre lesquels s'exercent des forces d'attraction. Les ions se forment par suite du transfert d'électrons d'un atome à un autre. En règle générale, les enveloppes électroniques des ions correspondent aux enveloppes électroniques des gaz rares. La distribution des électrons à proximité de chaque ion est peu différente d'une distribution à symétrie sphérique et n'est que faiblement perturbée

dans les régions où se touchent les ions voisins. Comme exemples typiques de cristaux ioniques, on peut citer NaCl, LiF, KI, etc. Les cristaux ioniques sont caractérisés par de grandes *énergies de liaison*. On entend par énergie de liaison le travail qu'il faudrait fournir pour décomposer le système, à la température du zéro absolu, en ses constituants. Dans le cas des cristaux ioniques, ces constituants sont des ions positifs et négatifs dont est construit le réseau. L'énergie de liaison du réseau de NaCl est égale à 764 kJ/mole (183 kcal/mole). Cette énergie est près de 1,7 fois plus grande que l'énergie de liaison des ions Na^+ et Cl^- dans la molécule NaCl, en admettant que dans la molécule ces ions se trouvent à la même distance l'un de l'autre que les ions voisins dans le réseau cristallin. L'énergie de cohésion des réseaux de cristaux ioniques analogues est du même ordre de grandeur. Les cristaux ioniques sont mauvais conducteurs d'électricité à basse température, mais présentent une grande conductibilité aux températures élevées.

3. Dans les cristaux covalents (diamant, silicium, germanium, etc.) la liaison est *homéopolaire* (comme dans la molécule d'hydrogène), ce qui signifie qu'elle est assurée par les *forces d'échange*. Les cristaux covalents possèdent généralement des réseaux cristallins où le nombre des plus proches voisins d'un atome donné est égal à la valence de l'élément chimique. Ces cristaux sont généralement formés par des éléments chimiques légers appartenant aux colonnes centrales du système périodique. Les électrons de valence des atomes formant le réseau sont *collectivisés*. Les cristaux covalents présentent une grande dureté, une faible conductibilité électrique à basse température lorsqu'ils sont purs. L'énergie de cohésion des réseaux des cristaux covalents est de l'ordre de 700 à 1200 kJ/mole (170 à 280 kcal/mole).

4. Les cristaux moléculaires sont des agrégats de molécules faiblement liés. La liaison est assurée par les *forces de Van der Waals*. Tous les cristaux organiques et nombre d'autres composés sont du type moléculaire. La liaison moléculaire est la seule liaison possible chez les cristaux formés par les atomes des gaz inertes. Pour des substances telles que Ar, CH_4, l'énergie de liaison est de l'ordre de 8 kJ/mole (2 kcal/mole). Les cristaux moléculaires ont de petites températures de fusion, d'ébullition et de sublimation et sont fortement compressibles.

5. Dans les métaux la liaison est assurée par les électrons se trouvant entre les ions du réseau cristallin; leur action est semblable à celle qu'ils assument dans la liaison homéopolaire entre les noyaux dans la molécule d'hydrogène. Il va de soi qu'il est impossible d'effectuer un calcul quantitatif précis dans un système à nombreuses particules comme l'est le métal. Aussi faut-il se contenter de considérations qualitatives. Sont métalliques les éléments dont les atomes ont des couches électroniques externes incomplètes. Etant assez

faiblement attachés aux noyaux atomiques, les électrons de ces couches peuvent passer d'un noyau à un autre. Même si l'électron rencontre une barrière de potentiel sur son trajet, il la surmonte aisément par effet tunnel. Les noyaux perdent ainsi leurs couches électroniques périphériques. Leurs électrons ne sont pas attachés à des atomes individuels, mais sont *collectivisés* et appartiennent au cristal tout entier. Ces électrons dits « libres » se comportent comme un gaz électronique. En vertu du principe de Pauli, les électrons ne peuvent rester sur place et sont animés d'un mouvement incessant, ce qui signifie que dans le métal le gaz électronique se trouve dans un *état de dégénérescence* (cf. t. II, § 82).

L'existence d'électrons libres détermine les grandes conductibilités électrique et thermique des métaux, leur éclat dit métallique, leurs propriétés mécaniques très particulières autorisant le forgeage et l'estampage.

Chacun des ions du réseau cristallin d'un métal est chargé positivement. De ce fait des forces de répulsion électriques s'exercent entre les ions. Les électrons libres équilibrent ces forces et retiennent les ions dans leurs positions d'équilibre. Chaque fois qu'un ion s'écarte de sa position d'équilibre, les électrons libres, étant légers et rapides, se redistribuent dans l'espace occupé par le métal de telle sorte qu'apparaissent des forces qui obligent l'ion à revenir à sa position d'équilibre. C'est ce qui assure la stabilité de l'édifice cristallin et du métal.

§ 56. Vibrations des atomes dans une chaîne unidimensionnelle rectiligne

1. Pour justifier d'une façon rigoureuse l'admissibilité et les limites d'application du modèle continu du corps solide adopté par Debye dans sa théorie de la capacité calorifique, il importe de considérer tout d'abord la question des vibrations du réseau cristallin en se plaçant à un point de vue atomistique conséquent, comme l'avaient fait Born et Karman. Examinons cette question en prenant pour exemple une chaîne unidimensionnelle rectiligne d'atomes. En procédant ainsi, on n'obtiendra assurément pas de résultats quantitatifs exacts, valables pour un corps tridimensionnel, mais on simplifiera grandement l'étude du problème tout en obtenant des résultats qualitatifs essentiels.

2. Supposons d'abord que tous les atomes de la chaîne sont identiques et que leurs positions d'équilibre sont équidistantes; soit a la distance de séparation des atomes. Prenons en considération uniquement les forces qu'exercent les deux atomes voisins sur un atome de la chaîne. L'action de tous les autres atomes est à négliger. Cette simplification est appelée *approximation des plus proches voisins.* Supposons que les atomes ne puissent subir que des déplacements

longitudinaux par rapport à leurs positions d'équilibre. Notons ξ_n le déplacement du n-ième atome. Le déplacement relatif $(\xi_n - \xi_{n-1})$ d'atomes voisins sera supposé petit devant le paramètre de réseau a. Lorsque le n-ième atome subit un déplacement par rapport au $(n-1)$-ième atome, apparaît la force $F_{n,\,n-1}$ qui s'exerce sur l'atome déplacé et qui est dirigée à l'encontre du déplacement relatif. Si les déplacements relatifs sont petits, on peut admettre que cette force est quasi élastique, i.e. $F_{n,n-1} = -\varkappa\,(\xi_n - \xi_{n-1})$, où le « coefficient d'élasticité » $\varkappa$ est une quantité constante pour la chaîne d'atomes considérée. La force totale s'exerçant sur l'atome est donc

$$F_{n,\,n-1} + F_{n,\,n+1} = -\varkappa\,(\xi_n - \xi_{n-1}) - \varkappa\,(\xi_n - \xi_{n+1}) = \varkappa\,(\xi_{n-1} - 2\xi_n + \xi_{n+1}),$$

et l'équation du mouvement est

$$m\ddot{\xi}_n = \varkappa\,(\xi_{n-1} - 2\xi_n + \xi_{n+1}), \qquad (56.1)$$

où m est la masse de l'atome.

Lorsque le nombre N d'atomes est très grand, il est extrêmement difficile de trouver une solution générale de l'équation (56.1). Pour trouver une solution particulière, considérons d'abord le cas où $N = \infty$, autrement dit le cas où la chaîne d'atomes se prolonge à l'infini des deux côtés. La chaîne d'atomes possède une *symétrie de translation*, c'est-à-dire qu'après un glissement d'un nombre entier de périodes a la chaîne se transforme en elle-même. On peut donc envisager l'existence d'une solution particulière de (56.1) correspondant à ce type de symétrie : tous les atomes exécutent des oscillations harmoniques identiques, mais leurs phases sont décalées d'une même quantité lorsqu'on passe d'un atome quelconque à l'atome voisin ayant un numéro plus grand. Cette solution se présente sous la forme d'une onde monochromatique progressive d'amplitude constante :

$$\xi = \xi_0 e^{i(\omega t - kx)}. \qquad (56.2)$$

Cette onde présente la particularité que l'argument x ne peut prendre que des valeurs discrètes $x_n = na$ $(n = -2, -1, 0, +1, +2, \ldots)$. Si on remplace le nombre d'onde k par $k = \frac{2\pi}{a} p$, où p est un nombre entier arbitraire, les oscillations des atomes de la chaîne n'en seront pas affectées. Par conséquent, sans perte de généralité, on peut limiter les variations de k par un intervalle de longueur $2\pi/a$ que l'on appelle *zone de Brillouin* (1889-1969). En particulier, l'intervalle

$$-\pi/a \leqslant k \leqslant \pi/a \qquad (56.3)$$

constitue la *principale zone de Brillouin*. Pour les valeurs de k positives, l'onde progresse en avant (à droite) et pour les k négatifs, l'onde régresse (se déplace à gauche). Pour ces valeurs de k la lon-

gueur d'onde Λ (qui est une grandeur essentiellement positive) peut varier dans les limites

$$\infty > \Lambda \geqslant 2a. \tag{56.4}$$

Par conséquent, la structure considérée étant discrète, il est inutile d'envisager la propagation d'ondes dont les longueurs sont inférieures à $2a$. Si on pose, par exemple, $\Lambda = a$, les déplacements de tous les atomes seront les mêmes à un instant donné, ce qui signifie que

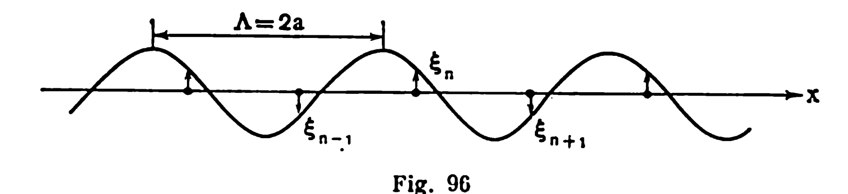


Fig. 96

toute la chaîne effectue un déplacement d'ensemble. Or cela équivaut à une longueur d'onde $\Lambda = \infty$ entrant dans l'intervalle (56.4).

Cherchons la condition pour que l'onde (56.2) soit une solution de l'équation (56.1). En remarquant que $\ddot{\xi}_n = -\omega^2\xi_n$ et en substituant cette expression dans (56.1), on constate que cette équation sera vérifiée si

$$\omega^2 = \frac{\varkappa}{m}(2 - e^{ika} - e^{-ika}) = 2\frac{\varkappa}{m}(1 - \cos ka).$$

En se limitant aux k positifs, on en tire

$$\omega = 2\sqrt{\frac{\varkappa}{m}} \sin\frac{ka}{2}, \tag{56.5}$$

vu que la fréquence ω est essentiellement positive.

3. La vitesse de phase de l'onde est

$$c = \frac{\omega}{k} = a\sqrt{\frac{\varkappa}{m}}\,\frac{\sin(ka/2)}{ka/2}, \tag{56.6}$$

autrement dit elle dépend de k et donc de Λ. Il en résulte qu'il y a *dispersion* et c'est ce qui explique que la formule (56.5) est appelée *formule de dispersion*. La vitesse de groupe de l'onde est égale à

$$u = \frac{d\omega}{dk} = a\sqrt{\frac{\varkappa}{m}} \cos\frac{ka}{2}. \tag{56.7}$$

Pour $ka = \pi$ (i.e. pour $\Lambda = 2a$) u devient égal à zéro. Dans ce cas, *l'onde ne transporte pas d'énergie*. L'examen de la figure 96 laisse apparaître la cause physique de cette situation. Sur cette figure les fléchettes représentent les déplacements instantanés des atomes pour $\Lambda = 2a$. On voit que dans ce cas $\xi_{n-1} = \xi_{n+1} = -\xi_n$, de sorte que

l'équation (56.1) prend la forme

$$m\ddot{\xi}_n = -4\varkappa\xi_n.$$

Cette nouvelle équation montre que dans le cas considéré les atomes se comportent comme s'ils n'étaient pas couplés, chacun d'eux étant isolé et effectuant des oscillations harmoniques avec la fréquence $\omega = 2\sqrt{\varkappa/m}$. Les oscillations des atomes voisins sont déphasées de π. Comme les forces d'interaction et les déplacements de n'importe

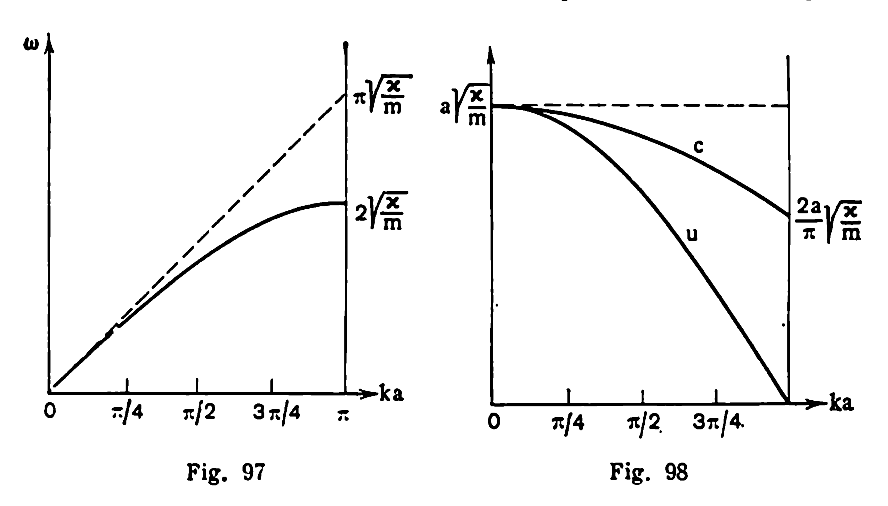


Fig. 97 Fig. 98

quelle paire d'atomes voisins sont, à un instant donné, égaux et opposés, le travail qu'effectue l'atome 1 sur l'atome 2 est exactement égal au travail qu'effectue l'atome 2 sur l'atome 1. Cela signifie que dans l'approximation des plus proches voisins il n'y a aucun transport d'énergie d'un atome à un autre. En ce qui concerne la vitesse de phase c, pour $ka = \pi$ elle est égale à

$$c = \omega/k = 2\sqrt{\varkappa/m} : (\pi/a) = 2\frac{a}{\pi}\sqrt{\frac{\varkappa}{m}}.$$

La dépendance de la fréquence angulaire ω, de la vitesse de phase et de la vitesse de groupe avec le nombre d'onde k est représentée sur les figures 97 et 98. Pour k petits, i.e. aux grandes longueurs d'onde, la formule (56.5) se réduit à $\omega = ka\sqrt{\varkappa/m} =$ const, et les deux vitesses limites dégénèrent en $c = u = a\sqrt{\varkappa/m} =$ const. Dans ce cas il n'y a plus de dispersion et la chaîne d'atomes se comporte comme un milieu continu. Dans ces conditions, on peut déterminer la valeur de la vitesse $c = u$ de la formule établie en théorie de l'élasticité pour la vitesse du son dans une barre (cf. t. I, § 81). En effet, le module d'Young E de la chaîne d'atomes doit être calculé par la

formule $F_{n,n-1} = E(\xi_n - \xi_{n-1})/a$, puisque $(\xi_n - \xi_{n-1})/a$ est la dilatation relative de la chaîne. Or, comme la force de tension est égale à $F_{n,n-1} = \varkappa(\xi_n - \xi_{n-1})$, il s'ensuit que $E = \varkappa a$. Le rôle de la densité est assumé ici par $\rho = m/a$. C'est pourquoi la formule de la vitesse du son dans une barre entraîne

$$u = c = \sqrt{E/\rho} = a\sqrt{\varkappa/m},$$

ce qui est un résultat correct. Par conséquent, dans le cas d'ondes longues ($\Lambda \gg a$), on peut calculer la fréquence des oscillations par les formules établies pour le modèle continu de la chaîne d'atomes. Même dans le cas des ondes les plus courtes ($\Lambda = 2a$), l'erreur que l'on commet en procédant ainsi n'est pas très grande : on trouve pour ω la valeur $\pi\sqrt{\varkappa/m}$, c'est-à-dire une valeur près de 1,5 fois plus grande que la valeur exacte $2\sqrt{\varkappa/m}$.

4. Dans le cas général, on peut représenter le déplacement de la chaîne d'atomes comme une superposition d'ondes de fréquences différentes se propageant en avant et en arrière. Certes, toute chaîne réelle est limitée ; en notant n le nombre d'atomes qui la composent, sa longueur sera égale à $l = (n - 1)a$. Si les atomes ne peuvent exécuter que des oscillations longitudinales, le nombre de degrés de liberté de la chaîne sera égal à n. Si nous fixons les atomes extrêmes, le nombre de degrés de liberté diminuera de deux et sera égal à $n - 2$. L'équation (56.1) aura pour solution particulière l'onde (56.2) se propageant en avant ($k \geqslant 0$). Une onde de même fréquence mais se propageant en arrière sera aussi solution de la même équation. Par conséquent, la superposition de ces deux ondes sera aussi une solution particulière de l'équation

$$\xi = \xi_0 e^{i(\omega t - kx)} + \eta_0 e^{i(\omega t + kx)} = e^{i\omega t}(\xi_0 e^{-ikx} + \eta_0 e^{ikx}).$$

Mais comme les atomes extrêmes sont fixés, leurs déplacements doivent être égaux à zéro à tout instant. On doit donc avoir

$$\xi_0 + \eta_0 = 0, \quad \xi_0 e^{-ikl} + \eta_0 e^{ikl} = 0.$$

De la première relation il vient $\eta_0 = -\xi_0$. Compte tenu de ce résultat, la seconde relation donne

$$e^{ikl} - e^{-ikl} = 0,$$

ou

$$\sin kl = 0. \tag{56.8}$$

On a donc obtenu une onde stationnaire

$$\xi = \xi_0 e^{i\omega t}(e^{-ikx} - e^{ikx}) = -2i\xi_0 \sin kx\, e^{i\omega t} =$$
$$= 2\xi_0 \sin kx \sin \omega t, \tag{56.9}$$

avec

$$kl = N\pi \quad (N = 0,\ 1,\ 2,\ \ldots,\ N_{max}), \tag{56.10}$$

ou encore

$$l = N\frac{\Lambda}{2},$$

i.e. *la chaîne doit contenir un nombre entier de demi-longueurs d'ondes.* A chaque onde stationnaire (56.9) correspond une vibration propre (ou normale) de la chaîne.

La valeur maximale de N correspond à la valeur maximale de $k = \pi/a$, i.e. $N_{max} = l/a = n - 1$. Néanmoins, on doit exclure cette valeur tout comme la valeur $N = 0$, puisqu'à ces valeurs correspondent des valeurs de k pour lesquelles tous les atomes se déplacent également. Or, comme les atomes extrêmes sont fixés, tous les autres atomes doivent être immobiles. Il s'ensuit que dans une chaîne à extrémités fixes ne peuvent s'exciter que $(n - 2)$ vibrations normales, donc un nombre exactement égal au nombre de degrés de liberté de la chaîne. Le mouvement d'ensemble d'une chaîne à extrémités fixes peut être représenté par la superposition de ces $(n - 2)$ vibrations normales.

On arrive à la conclusion que *pour le cas des ondes longues on peut utiliser le modèle continu des corps solides.* C'est ce qui est mis à profit dans la théorie de Debye pour calculer la capacité calorifique des corps solides aux basses températures lorsque seules les ondes longues possèdent une énergie appréciable. On arrive à la même conclusion dans le cas général en analysant les vibrations d'un réseau cristallin tridimensionnel composé de particules identiques (corps solide monoatomique). Mais dans un corps tridimensionnel, même en négligeant l'anisotropie, aux *vibrations longitudinales*, considérées ci-dessus pour la chaîne unidimensionnelle, viennent s'ajouter encore deux branches de *vibrations transversales*, i.e. des vibrations orthogonales au vecteur d'onde k et s'effectuant dans des plans orthogonaux. Le nombre de vibrations normales, ainsi que le nombre de degrés de liberté, est dans ce cas multiplié par trois.

5. Jusqu'à présent on avait supposé que le cristal était composé d'atomes *identiques*. Dans le cas de cristaux dont la maille élémentaire contient plusieurs atomes *différents*, on doit tenir compte des vibrations relatives de ces différents atomes. Pour faire une étude qualitative de ce cas, revenons à notre modèle de chaîne unidimensionnelle mais composée cette fois de deux atomes différents alternant dans la chaîne (fig. 99). Notons M et m $(M > m)$ les masses de ces atomes, et ξ_n et η_n leurs déplacements par rapport à leurs positions d'équilibre; comme ci-dessus les calculs seront faits dans l'approximation des plus proches voisins, ce qui revient à ne tenir compte que des forces d'interaction s'exerçant entre atomes voisins. La figure 99 représente la disposition des atomes dans la chaîne occupant leurs positions d'équilibre. Au-dessous de chaque atome est indiqué son déplacement par rapport à sa position d'équilibre. En se fondant sur

cette figure, on obtient, par analogie avec l'équation (56.1), le résultat suivant :

$$\begin{aligned} M\ddot{\xi}_n &= \varkappa\,(\eta_{n-1} - 2\xi_n + \eta_n), \\ m\ddot{\eta}_n &= \varkappa\,(\xi_{n+1} - 2\eta_n + \xi_n). \end{aligned} \tag{56.11}$$

On cherche ensuite une solution particulière de ce système d'équations sous forme d'une onde monochromatique progressive

$$\xi_n = \xi_0 e^{i(\omega t - kna)}, \quad \eta_n = \eta_0 e^{i(\omega t - kna)}, \tag{56.12}$$

en supposant que a varie dans l'intervalle (56.3), en entendant main-

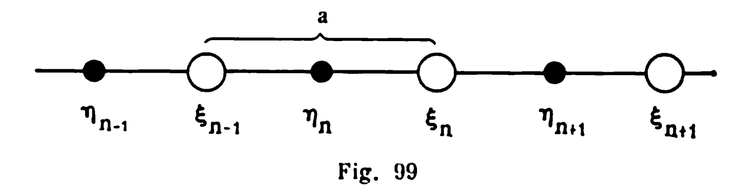


Fig. 99

tenant par a la distance entre deux atomes voisins de même sorte. Après substitution dans (56.11) on obtient

$$\begin{aligned} (M\omega^2 - 2\varkappa)\,\xi_0 + \varkappa\,(1 + e^{ika})\,\eta_0 &= 0, \\ \varkappa\,(1 + e^{-ika})\,\xi_0 + (m\omega^2 - 2\varkappa)\,\eta_0 &= 0. \end{aligned} \tag{56.13}$$

Après élimination de ξ_0 et η_0 on trouve

$$Mm\omega^4 - 2\varkappa\,(M + m)\,\omega^2 + 2\varkappa^2\,(1 - \cos ka) = 0,$$

d'où

$$\omega^2 = \frac{\varkappa}{Mm}\left(M + m \pm \sqrt{M^2 + m^2 + 2Mm\cos ka}\right). \tag{56.14}$$

C'est cette formule qui décrit le spectre des fréquences propres des vibrations de la chaîne d'atomes. Comme on trouve les signes « plus » et « moins » devant la racine carrée, on obtient deux branches de fréquences. Au signe « moins » correspond la fréquence $\omega_1 = \omega_1(k)$ et au signe « plus » la fréquence $\omega_2 = \omega_2(k)$. La branche $\omega_1 = \omega_1(k)$ est appelée *branche acoustique* ou *branche de Debye*, et la branche $\omega_2 = \omega_2(k)$ est la *branche optique* ou *branche de Born*; elles sont représentées sur la figure 100. Lorsque k est petit (ondes longues), la fréquence ω est petite et varie linéairement avec k. D'après (56.13) on a alors $\xi_0 = \eta_0$ et par suite $\xi_n = \eta_n$; cela signifie que les atomes voisins de masses M et m (ainsi que tous les atomes se trouvant sur un segment petit devant la longueur d'onde) vibrent *en phase*. Dans ce cas on peut approximer la chaîne d'atomes par un modèle unidimensionnel continu.

La branche $\omega_2 = \omega_2(k)$ présente la particularité que pour $k \to 0$ ω_2, au lieu de tendre vers zéro, tend vers sa valeur maximale. Pour k petit chacune des équations (56.13) se réduit à $M\xi_0 = -m\eta_0$ et par suite $M\xi = -m\eta$. Cela signifie que dans ce cas les atomes voisins de masses M et m vibrent *en opposition de phase*, i.e. un atome oscille par rapport à son voisin.

6. Dans un réseau cristallin tridimensionnel dont la maille élémentaire contient s atomes, il existe $3s$ branches de *vibrations normales*. Parmi ces branches trois sont des *branches acoustiques* correspondant l'une à des vibrations *longitudinales* et les deux autres à des vibrations *transversales*. Les $3s - 3$ branches restantes sont des *branches*

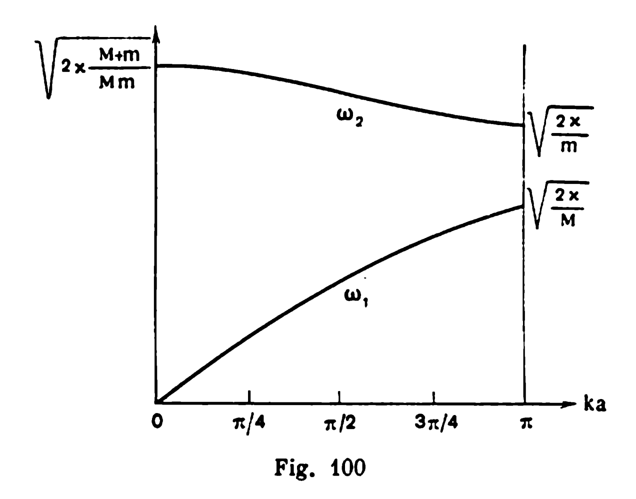


Fig. 100

optiques. Les fréquences de certaines vibrations normales peuvent coïncider en raison de la symétrie du réseau. Pour les ondes longues, les fréquences des vibrations acoustiques tendent vers zéro et sont proportionnelles au vecteur d'onde. Dans ce cas les atomes voisins de la maille élémentaire vibrent *en phase* et le cristal peut être assimilé à un *milieu continu*. Les vibrations optiques ont des fréquences élevées qui ne s'annulent pas pour les ondes de longueur infinie. Lors des vibrations optiques les atomes de la maille élémentaire subissent des déplacements importants les uns par rapport aux autres.

La diffusion de la lumière par les ondes acoustiques thermiques donne lieu à une variation de la fréquence, ce qui constitue l'*effet Mandelstam-Brillouin* (cf. t. IV, § 99, ainsi que le problème n° 2 au § 57). La branche optique des vibrations est caractérisée par les fréquences $\nu = \omega/2\pi \sim 10^{12}$ à 10^{13} Hz qui se situent dans la région infrarouge du spectre, ce qui explique leur dénomination. Les vibrations optiques peuvent donner lieu à une variation des moments électriques de la maille élémentaire (c'est ce qui se produit dans les

cristaux de NaCl lorsque les ions Na^- et Cl^+ se déplacent l'un par rapport à l'autre). Dans ce cas on voit apparaître des bandes d'absorption dans l'infrarouge et une dispersion anomale correspondante dans le spectre optique du cristal. C'est aux vibrations de fréquences infrarouges qu'est liée la *diffusion Raman* (cf. t. IV, § 100).

Il est évident que par suite du caractère discret du réseau spatial le vecteur d'onde $\boldsymbol{k}$ ne peut être défini de façon univoque (comme dans le cas du réseau unidimensionnel). On peut lui ajouter un vecteur arbitraire $2\pi \boldsymbol{K}$ afin de modifier la phase des vibrations $\omega t - \boldsymbol{kr}$ de tous les atomes de $2\pi n$, n étant un nombre entier quelconque (positif ou négatif). Physiquement ce changement de la phase des vibrations ne se manifeste d'aucune manière. Si $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ sont les vecteurs de base du réseau cristallin, le vecteur $\boldsymbol{K}$ est défini par l'expression

$$\boldsymbol{K} = n_1\boldsymbol{a}_1^* + n_2\boldsymbol{a}_2^* + n_3\boldsymbol{a}_3^*, \tag{56.15}$$

où n_1, n_2, n_3 sont des nombres entiers arbitraires (positifs et négatifs) et $\boldsymbol{a}_1^*$, $\boldsymbol{a}_2^*$, $\boldsymbol{a}_3^*$ les vecteurs de base du réseau réciproque:

$$\boldsymbol{a}_1^* = \frac{[\boldsymbol{a}_2\boldsymbol{a}_3]}{(\boldsymbol{a}_1[\boldsymbol{a}_2\boldsymbol{a}_3])}, \quad \boldsymbol{a}_2^* = \frac{[\boldsymbol{a}_3\boldsymbol{a}_1]}{(\boldsymbol{a}_2[\boldsymbol{a}_3\boldsymbol{a}_1])}, \quad \boldsymbol{a}_3^* = \frac{[\boldsymbol{a}_1\boldsymbol{a}_2]}{(\boldsymbol{a}_3[\boldsymbol{a}_1\boldsymbol{a}_2])}. \tag{56.16}$$

(voir t. II, § 130 et t. I, § 7). Le plus petit domaine de variation du vecteur $\boldsymbol{k}$ qui peut suffire pour représenter n'importe quelles vibrations des atomes du réseau est appelé *zone de Brillouin*. On peut notamment faire en sorte que le point $\boldsymbol{k} = 0$ dans l'espace des vecteurs d'onde soit un point de symétrie par réflexion. Dans ce cas la zone de Brillouin est dite *fondamentale*. Le vecteur $\boldsymbol{K}$ (56.15) correspondant à des nombres entiers n_1, n_2, n_3 arbitraires est appelé *vecteur du réseau réciproque*.

PROBLÈME

Expliquer pourquoi la formule (56.14) ne se transforme pas en formule (56.5) lorsque $M = m$?

§ 57. Les phonons et les quasi-particules



1. Le mouvement interne d'un corps au repos peut être décrit si on définit les mouvements de toutes les particules individuelles composant ce corps. Ce procédé pourrait être appelé description *microscopique* (ou *individuelle*) du mouvement. On peut utiliser aussi un procédé de description *collectif* dans lequel le mouvement du corps tout entier est considéré comme le résultat de la superposition de tous les mouvements auxquels participent toutes les particules du corps. Ce deuxième procédé peut être préférable lorsque toutes les particules du corps interagissent entre elles. Dans ce cas la décomposition du mouvement total du corps en mouvements collectifs

constitutifs doit être effectuée de telle sorte que chacun de ces mouvements collectifs puisse être excité individuellement. Si le mouvement excité du corps n'est pas trop intense, il peut toujours être décomposé en ondes monochromatiques planes de fréquences différentes se propageant dans le corps dans des directions différentes de façon pratiquement indépendante les unes des autres. Lorsque l'excitation du corps s'intensifie, on voit apparaître des effets *non linéaires*. Cependant si les écarts à la linéarité ne sont pas trop notables, on peut encore utiliser la décomposition en ondes monochromatiques planes, quoique apparaissent alors des interactions entre ces ondes.

En physique classique, les deux procédés de description du mouvement sont en principe équivalents. En mécanique quantique, on donne la préférence au deuxième procédé de description et ce, pour des raisons de quantification. Dans sa théorie de la capacité calorifique (cf. § 54), Debye effectua déjà avec succès la quantification de l'énergie des ondes monochromatiques stationnaires auxquelles se laisse réduire le mouvement du corps. En théorie de la capacité calorifique, il n'est pas nécessaire de décomposer les ondes stationnaires en ondes progressives puisque dans ce cas on s'intéresse à l'énergie du corps se trouvant dans un état d'équilibre statistique et non pas à son impulsion qui, pour un corps au repos, est égale à zéro à tout instant. Mais dans l'étude des différents processus évoluant dans les corps, on doit procéder à la décomposition du mouvement en ondes progressives et à leur quantification même dans le cas d'un équilibre statistique local.

Conformément à l'hypothèse de de Broglie, qui est confirmée par des données expérimentales, à chaque onde monochromatique progressive sont liées une énergie et une impulsion définies par les relations

$$\mathscr{E} = \hbar\omega, \quad p = \hbar k, \tag{57.1}$$

qui ont été introduites par analogie avec la théorie des photons. Une onde transportant une énergie et une impulsion définies par les formules (57.1) se comporte à un certain point de vue comme une particule. Une particule assimilable en ce sens à une onde acoustique est appelée *phonon*. On ne doit pas attribuer au concept de phonon autre chose que ce que comporte sa définition. Le phonon transporte une énergie et une impulsion qui sont liées à la fréquence ω de l'onde et au nombre d'onde k par l'intermédiaire de la constante de Planck $\hbar$. Il est par contre dénué de sens de parler de la forme et des dimensions d'un phonon en l'imaginant comme une petite bille. C'est pour cela qu'en parlant de phonons on les appelle des « quasi-particules » et non des particules, et leurs impulsions sont appelées « quasi-impulsions ». Les considérations qu'on invoque plus loin justifieront pleinement cette terminologie.

Les grandeurs ω et k ne possèdent des valeurs bien déterminées que dans le cas d'ondes planes illimitées, mais n'existent réellement que les ondes *spatialement limitées.* Une image réelle du phonon est une onde limitée, par exemple un paquet d'ondes, et non pas une onde illimitée.

2. L'hypothèse des phonons est conforme notamment à l'existence de la *pression acoustique.* Considérons, pour simplifier, un milieu continu isotrope (un corps solide isotrope, un liquide ou un gaz) dans lequel se propage une onde acoustique longitudinale monochromatique plane, tombant sous incidence normale sur une surface plane d'un corps solide qui l'absorbe. Cette onde transmet par seconde, à l'unité de surface du corps absorbant, l'impulsion $cN\hbar k$, où c est la vitesse du son et N le nombre de phonons contenus dans l'unité de volume du milieu. Cette impulsion est précisément la pression $\mathcal{P}$ qu'exerce le son sur le corps. Comme le milieu continu n'est pas dispersif, la vitesse du son c s'y confond avec la vitesse de phase $c_{\text{phase}} = \omega/k$. Par suite

$$\mathcal{P} = \varepsilon, \tag{57.2}$$

où $\varepsilon = N\hbar\omega$ est la densité volumique de l'énergie acoustique tombant sur le corps. La formule (57.2) reste valable dans le cas général d'une incidence normale de l'onde même si elle subit une réflexion et une transmission, à cela près que la densité d'énergie est alors donnée par l'expression $\varepsilon = (1 + r)\, N\hbar\omega$, où r est le *coefficient de réflexion.* Les résultats obtenus sont conformes à l'expérience et aux résultats de l'hydrodynamique classique.

3. Dans les corps solides continus et isotropes deux types de phonons peuvent être excités : les *phonons longitudinaux* et les *phonons transversaux.* Dans le cas de milieux isotropes, les fréquences des phonons transversaux ne dépendent que de la longueur d'onde et sont indépendantes de leur polarisation. Outre les phonons transversaux et longitudinaux, il peut apparaître dans les cristaux des phonons d'autres types correspondant à des fréquences et à des polarisations différentes des vibrations. (En général, dans les cristaux, les ondes « longitudinales » et « transversales » ne sont pas rigoureusement longitudinales ou transversales.) Les relations (57.1) s'appliquent à tous ces phonons, mais la fréquence ω des vibrations n'est pas nécessairement liée au vecteur d'onde $\boldsymbol{k}$ par une relation homogène linéaire, comme c'est le cas lorsqu'il n'y a pas de dispersion ; les relations $\omega = \omega(\boldsymbol{k})$ sont alors plus compliquées et on les appelle *relations de dispersion.* Elles sont différentes pour des phonons de types différents. Comme exemples de relations de dispersion, on peut citer les formules (56.5) et (56.14) qui ont été établies pour des chaînes d'atomes unidimensionnelles.

Dans un réseau cristallin le vecteur d'onde $\boldsymbol{k}$ est défini de façon *non univoque*, mais à un terme près qui est égal au vecteur du réseau

réciproque (cf. § 56, 6). En conséquence, le vecteur $p = \hbar k$ est également défini de façon *non univoque*. On peut cependant éliminer cet inconvénient en limitant le domaine de variation de k par la zone de Brillouin fondamentale; c'est ce qu'on fera dans ce qui suit. Ainsi, le vecteur p univoquement défini est appelé *quasi-impulsion* du phonon. Le phonon lui-même, parce que jouissant des propriétés de particule, est appelé *quasi-particule*. On a introduit ce terme pour bien indiquer qu'il ne s'agit pas d'une vraie particule.

On introduit la notion de quasi-particule dans d'autres parties de la physique, par exemple les photons existant dans le vide et surtout dans les milieux sont des quasi-particules. Cette conception des photons est corroborée par des phénomènes tels que les interférences, l'effet Doppler, l'effet Vavilov-Cerenkov, la variation de la fréquence de la lumière se propageant dans un champ de gravitation. On a déjà examiné ces différents phénomènes en postulant l'existence de quasi-particules de lumière sans utiliser nommément ce terme.

4. Dans un réseau cristallin parfait ne comportant ni impuretés ni défauts (atomes d'impuretés, atomes interstitiels, lacunes) une onde acoustique plane doit se propager dans l'approximation linéaire *sans subir d'amortissement* et *sans être diffusée de côté*. L'approximation linéaire consiste en ce que le développement de l'énergie potentielle du cristal en série de puissances des déplacements des atomes par rapport à leurs positions d'équilibre, est arrêté aux termes de *deuxième* degré. C'est alors qu'apparaissent les ondes avec les vibrations harmoniques des atomes ou phonons qui *n'interagissent pas entre elles*. Si le développement comporte des termes d'ordres supérieurs et s'ils sont suffisamment petits (en règle générale, ils sont petits tant que le réseau subsiste, i.e. jusqu'à sa température de fusion), on peut également parler de la propagation d'ondes planes (ou phonons) dans le cristal. Mais dans ce cas *les différentes ondes ou phonons entrent en interaction*. Comme l'énergie et l'impulsion des phonons sont quantifiées, l'interaction des phonons présente le caractère de *chocs* lors desquels les phonons initiaux sont détruits et de nouveaux phonons sont créés. La présence dans l'expression de l'énergie potentielle de termes de troisième degré détermine l'éventualité de chocs de trois phonons simultanément, et si en plus il y a des termes de quatrième degré, on doit envisager l'éventualité de chocs de quatre phonons simultanément, etc.

Les phonons et, d'une manière générale, toutes les quasi-particules, conviennent particulièrement bien à la description d'excitations collectives *faibles* dans les corps solides. Entre deux chocs successifs le phonon se déplace librement et si sa « longueur de libre parcours » est assez grande par rapport au paramètre du réseau cristallin, il est permis d'assimiler l'état excité du cristal à un *gaz de phonons*; comme le nombre de phonons ne se conserve pas, on peut les considérer comme des *bosons*.

La figure 101 donne, sous forme graphique, quelques exemples d'interactions des phonons. Les phonons y sont figurés par des fléchettes et les processus de leurs interactions par de petits cercles. La figure 101, *a* illustre le processus de désintégration du phonon $\boldsymbol{k}$, ω en deux phonons $\boldsymbol{k}_1$, ω_1 et $\boldsymbol{k}_2$, ω_2, et la figure 101, *b* illustre la fusion

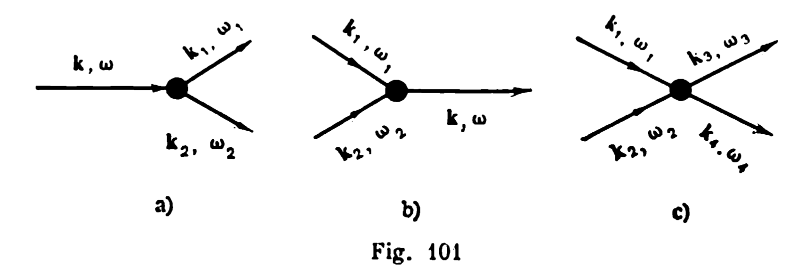


Fig. 101

de deux phonons $\boldsymbol{k}_1$, ω_1 et $\boldsymbol{k}_2$, ω_2 en un seul phonon $\boldsymbol{k}$, ω. Enfin, la figure 101, *c* représente le choc de deux phonons $\boldsymbol{k}_1$, ω_1 et $\boldsymbol{k}_2$, ω_2 qui a pour résultat l'apparition de deux nouveaux phonons $\boldsymbol{k}_3$, ω_3 et $\boldsymbol{k}_4$, ω_4.

Les interactions des phonons vérifient la *loi de la conservation de l'énergie*. Pour le processus illustré par la figure 101, *a*, la loi de la conservation de l'énergie s'exprime par

$$\hbar\omega = \hbar\omega_1 + \hbar\omega_2. \qquad (57.3)$$

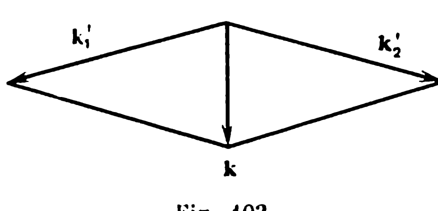


Fig. 102

Dans tous les autres cas on trouve des relations analogues. La loi de la conservation de la quasi-impulsion *n'est pas nécessairement vérifiée* en raison de la définition non univoque du vecteur d'onde $\boldsymbol{k}$, chose qui a déjà été signalée dans ce qui précède. Pour bien le montrer, décomposons le vecteur $\boldsymbol{k}$ en deux vecteurs: $\boldsymbol{k} = \boldsymbol{k}'_1 + \boldsymbol{k}'_2$ (fig. 102). En supposant que le vecteur $\boldsymbol{k}$ appartient à la zone de Brillouin fondamentale, selon la définition donnée plus haut, $\hbar\boldsymbol{k}$ est la quasi-impulsion. Si nous supposons que les vecteurs $\boldsymbol{k}'_1$ et $\boldsymbol{k}'_2$ ou au moins l'un de ces vecteurs sont trop longs pour être contenus en entier dans la première zone de Brillouin, d'après la même définition les vecteurs $\hbar\boldsymbol{k}'_1$ et $\hbar\boldsymbol{k}'_2$ ne seront plus des quasi-impulsions; pour en déduire les quasi-impulsions $\hbar\boldsymbol{k}_1$ et $\hbar\boldsymbol{k}_2$ il faut leur ajouter des vecteurs de la forme $2\pi\boldsymbol{K}$, où $\boldsymbol{K}$ est le vecteur du réseau réciproque (56.15). Cela signifie que pour le processus (fig. 101, *a*), par exemple, on devra écrire

$$\hbar\boldsymbol{k}_1 + \hbar\boldsymbol{k}_2 = \hbar\boldsymbol{k} + 2\pi\boldsymbol{K}. \qquad (57.4)$$

Lorsque $K = 0$, la quasi-impulsion se conserve dans le processus d'interaction des phonons. Ces processus sont dits *normaux*. Si $K \neq 0$, les processus correspondants sont dits *processus de rejet*. L'existence de processus de rejet justifie une nouvelle fois l'introduction des termes « quasi-particule » et « quasi-impulsion » au lieu des termes « particule » et « impulsion ».

Il va de soi que les relations de la forme (57.3) et (57.4) s'appliquent non seulement aux cas d'interaction mutuelle des phonons, mais également à tous les cas d'interaction des phonons avec d'autres particules ou quasi-particules, avec les photons par exemple. En termes de la physique classique, ces relations expriment les *lois d'interférences des ondes*, le *principe de Doppler* et, d'une façon générale, les *lois de variation de la fréquence des ondes dans différents processus*. C'est pour cela que la diffusion combinatoire de la lumière, la diffusion de Mandelstam-Brillouin, l'effet Vavilov-Cerenkov, la variation de la fréquence de la lumière se propageant dans un champ de gravitation et différents autres phénomènes peuvent être interprétés aussi bien en termes de la théorie ondulatoire qu'en utilisant le concept de quasi-particules.

5. Utilisons le concept de phonons pour faire une étude de la *conductibilité thermique des corps solides* ayant en vue les corps diélectriques et non pas les métaux. Dans les diélectriques, le transport de la chaleur est assuré par les *phonons*, tandis que dans les métaux le rôle principal revient, dans ce processus, aux *électrons*. Le concept même de conductibilité thermique concerne un état *d'équilibre local* d'un corps immobile, à chaque point duquel on peut attribuer une température déterminée. Pour amener le corps dans un état d'équilibre complet, on peut l'enfermer, par exemple, dans une enveloppe rigide maintenue à température constante. Dans ces conditions, par suite de l'excitation thermique des phonons, de leur absorption et de leur diffusion par d'autres phonons, par les impuretés et les défauts du réseau, par les surfaces délimitant le corps et par l'enveloppe qui le renferme, le corps parviendra finalement à un *état d'équilibre thermodynamique complet*, univoquement défini par la seule température de l'enveloppe. Cet état d'équilibre est caractérisé par une valeur bien déterminée de la densité spatiale de l'énergie des phonons, par une distribution déterminée de l'énergie dans le spectre des fréquences et par une distribution chaotique (dans le cas particulier, par une distribution isotrope) des directions de propagation des phonons. A ce point de vue l'état d'équilibre des phonons contenus dans une cavité rappelle l'état analogue des photons, i.e. le rayonnement du corps noir. L'état d'équilibre local d'un corps se distingue de l'état d'équilibre complet par ce que la température du corps varie d'un point à l'autre, et tous les autres paramètres caractérisant l'état d'un corps contenant des phonons arrivent à atteindre pratiquement les valeurs d'équilibre correspondant à cette

température. Dans les études de la conductibilité thermique, il s'agit toujours d'états d'équilibre local.

6. Supposons d'abord qu'on a affaire à un cristal parfait ne contenant aucune impureté chimique, le réseau cristallin ne comportant aucun défaut. On assimilera l'ensemble des phonons contenus dans le corps à un *gaz phononique* et on caractérisera sa conductibilité thermique par la formule

$$\varkappa \approx \frac{1}{3} \bar{v} C \lambda \tag{57.5}$$

établie dans le cadre de la théorie élémentaire des gaz (cf. t. II, § 89). Ici C est la capacité calorifique de l'unité de volume du corps (en notations du t. II, $C = nmc_v$), $\bar{v}$ la vitesse moyenne des phonons dans le corps et λ la longueur du libre parcours moyen des phonons. La quantité $\bar{v}$ est assimilée à la vitesse moyenne du son dans le corps, C est défini dans la théorie quantique de la capacité calorifique des corps solides. Ces deux quantités peuvent être mesurées expérimentalement. C'est la détermination de λ qui présente de grosses difficultés. Dans l'approximation harmonique (linéaire) les ondes acoustiques (les phonons) se propagent dans un cristal parfait sans y rencontrer aucun obstacle. Dans cette approximation, il n'y a pas de chocs mutuels entre les phonons. Si le cristal était illimité, aussi bien λ que la conductibilité thermique $\varkappa$ seraient infiniment grands. Dans les approximations d'ordres supérieurs, lorsqu'on tient compte, dans le développement de l'énergie potentielle du réseau, des termes du troisième degré et des degrés supérieurs par rapport aux déplacements des atomes de leurs positions d'équilibre, on voit apparaître les chocs mutuels des phonons qui limitent leurs longueurs de libre parcours. Ce sont les termes de troisième degré déterminant les chocs mutuels des trois phonons qui sont les plus importants (voir fig. 101, *a* et *b*).

7. Toutefois, *les chocs normaux entre les phonons n'affectent pas la conductibilité thermique du cristal.* La situation est ici la même que dans le cas du passage d'un courant électrique à travers un métal : la conductibilité électrique des métaux est liée à la longueur de libre parcours des électrons et des trous, mais les chocs mutuels entre les électrons et les trous ne l'affectent nullement (cf. t. III, § 42, 2). En effet, dans les chocs normaux, la quasi-impulsion du phonon coïncide avec l'impulsion vraie, et celle-ci se conserve lors des chocs, de même d'ailleurs que l'énergie. Lors de la désintégration, le phonon disparaît en donnant naissance à deux nouveaux phonons qui transportent alors la même énergie et la même impulsion à travers le cristal. Lors du choc de deux phonons, ceux-ci disparaissent en transmettant leurs énergies et leurs impulsions à deux nouveaux phonons qui les transportent à travers le cristal. La même chose se produit lors du choc simultané de trois phonons, ainsi que dans le cas où

quatre, cinq, etc., phonons entrent en collision. Par conséquent, les chocs mutuels normaux des phonons ne peuvent retarder la transmission de l'énergie ou de l'impulsion à travers le cristal. Si tous les chocs entre les phonons étaient normaux, la conductibilité thermique d'un réseau parfait illimité serait également infiniment grande.

On peut se demander alors pourquoi ces considérations ne s'appliquent pas à la conductibilité thermique d'un gaz composé de particules ordinaires (atomes et molécules), quoique dans ce cas il y a conservation de l'énergie et de l'impulsion lors des chocs entre ces particules. La raison en est que les chos mutuels des particules d'un gaz ordinaire ne s'accompagnent ni de leur désintégration ni de la naissance de nouvelles particules. Lors des chocs, la particule incidente ne se détruit pas et transmet son énergie et son impulsion à des particules déjà existantes et non pas à des particules nouvellement créées. En outre, dans le gaz, il n'y a pas de transport de substance et le transfert d'énergie n'est pas complet. A la suite d'un choc l'énergie de la particule incidente peut aussi bien diminuer qu'augmenter. Mais s'il existe un gradient de température dans le gaz, l'énergie des particules « chaudes » tend surtout à diminuer et l'énergie des particules « froides » à augmenter. C'est pour cela que s'établit dans le gaz un flux de chaleur orienté vers la région des plus basses températures.

Il résulte de ces considérations que *la conductibilité thermique d'un cristal parfait ne peut être liée qu'à des chocs de phonons qui donnent lieu à des processus de rejet*, puisque dans ce cas la loi de la conservation de la quasi-impulsion est en défaut. Il s'ensuit que pour calculer la longueur de libre parcours moyen des phonons figurant dans (57.5) on ne doit tenir compte que des chocs de ce type.

8. Il importe de remarquer que les calculs sont très laborieux et ne peuvent être effectués si on ne dispose pas de données précises concernant les forces d'interaction entre les atomes constituant le cristal ; or ces données manquent pour la majorité des cristaux. C'est pour cela que la théorie est loin d'être satisfaisante : on dispose de conceptions physiques claires, de procédés de calcul convenables mais qu'on ne peut mettre en œuvre pour obtenir des résultats quantitatifs définitifs. On se contentera donc de brèves considérations qualitatives.

A proximité du zéro absolu de température, lorsqu'il n'y a pratiquement pas de phonons thermiques, la longueur de libre parcours moyen des phonons est limitée par les dimensions du cristal. La situation ressemble à celle qui existe dans des gaz très raréfiés où la longueur de libre parcours des molécules est grande devant les dimensions du récipient contenant le gaz (cf. t. II, § 95). En posant dans la formule (57.5) $\lambda = l$, l étant les dimensions du cristal, on trouve une valeur de $\varkappa$ qui caractérisera non seulement les propriétés internes du cristal, mais dépendra encore de ses dimensions. A basse

température, la vitesse $\bar{v}$ est pratiquement constante et la capacité calorifique est, d'après la théorie de Debye, proportionnelle à T^3; par conséquent, la conductibilité thermique du cristal sera proportionnelle à T^3.

Lorsque la température s'élève, l'influence des dimensions du cristal passe au second plan et ce sont les *chocs mutuels des phonons* s'accompagnant de processus de rejet qui deviennent prédominants. De ce fait, ainsi que par suite de l'augmentation de la capacité calorifique, la conductibilité thermique augmente rapidement. Dans cette gamme de température, λ et donc la conductibilité thermique $\varkappa$ du cristal ne dépendent plus des dimensions du cristal et deviennent ses propriétés internes. Aux hautes températures, on peut s'attendre à une dépendance de la forme $\varkappa \sim \frac{1}{T}$. Effectivement, aux hautes températures la loi classique de l'équipartition de l'énergie cinétique suivant les degrés de liberté devient valable, ce qui implique que tous les phonons possèdent la même énergie (indépendante de la fréquence ω). Par suite, la densité N des phonons est proportionnelle à la densité d'énergie, donc à T, tandis que la capacité calorifique C atteint sa limite classique qui ne dépend pas de T. Par conséquent, la longueur de libre parcours moyen des phonons $\lambda \sim \frac{1}{N}$ et par suite la conductibilité thermique $\varkappa$ deviennent proportionnelles à $1/T$.

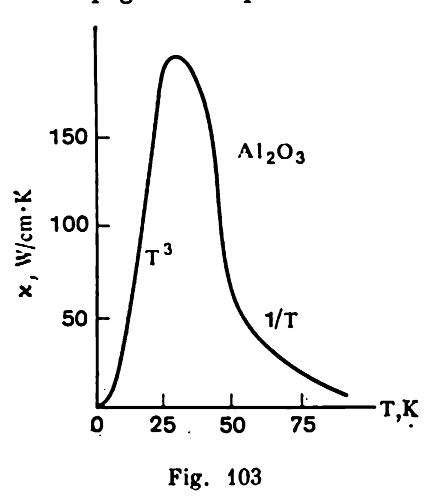


Fig. 103

La conclusion s'impose, de tout ce qui précède, que lors d'une élévation de température la conductibilité thermique d'un cristal diélectrique doit passer par un maximum. On le voit clairement sur la figure 103 qui représente la courbe expérimentale de la conductibilité thermique d'un cristal de saphir synthétique (Al_2O_3). Les maximums de $\varkappa$ ne sont pas toujours aussi nets avec les autres substances ; la cause en est la présence d'impuretés chimiques et de défauts du réseau cristallin qui réduisent sa conductibilité thermique.

PROBLÈMES

1. Calculer la pression exercée par une onde acoustique sur la surface de séparation de deux milieux continus, isotropes et non absorbants (fig. 104).

S o l u t i o n. Comme les composantes normales du flux d'énergie sont égales des deux côtés de la surface de séparation, on a

$$N_1 \varepsilon c_1 \cos \varphi - N_1' \varepsilon c_1 \cos \varphi = N_2 \varepsilon c_2 \cos \psi,$$

où N_1, N'_1 et N_2 sont les nombres de phonons incidents, réfléchis et transmis, contenus dans l'unité de volume, c_1 et c_2 les vitesses du son dans les milieux 1 et 2, $\varepsilon = \hbar\omega$ est l'énergie du phonon (la même dans les deux milieux). En introduisant le facteur de réflexion $r = N'_1/N_1$, on en tire

$$N_2 = \frac{c_1 \cos \varphi}{c_2 \cos \psi} (1 - r) N_1.$$

La pression $\mathscr{P}$ cherchée est égale à la composante normale de l'impulsion que transmet l'onde acoustique à l'unité de la surface de séparation des milieux:

$$\mathscr{P} = N_1 p_1 c_1 \cos^2 \varphi - (N_2 p_2 c_2 \cos^2 \psi - N'_1 p_1 c_1 \cos^2 \varphi).$$

Comme un milieu continu n'est pas dispersif, $pc = \varepsilon$. En utilisant les valeurs de N'_1 et N'_2 indiquées ci-dessus et en introduisant la densité d'énergie de l'onde acoustique incidente $\mathscr{E} = N_1 \varepsilon$, on obtient

$$\mathscr{P} = \mathscr{E} \cos \varphi \operatorname{cotg} \psi \, \{(1 + r) \cos \varphi \operatorname{tg} \psi - (1 - r) \sin \varphi\}. \quad (57.6)$$

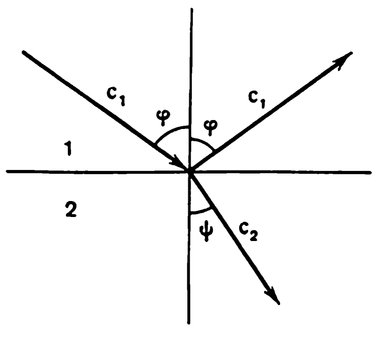


Fig. 104

2. Etablir la formule de la structure fine de Mandelstam-Brillouin (cf. t. IV, § 99) en se fondant sur le concept de phonons.

S o l u t i o n. Les relations entre l'énergie et l'impulsion pour la lumière (photon) et pour le son (phonon) se propageant dans un milieu sont de la forme

$$p_{\text{phot}} = \frac{n\mathscr{E}_{\text{phot}}}{c}, \qquad p_{\text{phon}} = \frac{\mathscr{E}_{\text{phon}}}{v_{\text{phon}}}, \quad (57.7)$$

où c est la vitesse de la lumière dans le vide et v_{phon} la vitesse du son (phonon) dans le milieu. La théorie classique conduit aux mêmes relations. Les équations de conservation de l'énergie et de l'impulsion lors de l'émission et de l'absorption d'un phonon sont

$$\mathscr{E}_{\text{son}} = \pm(\mathscr{E}_{\text{phot}} - \mathscr{E}'_{\text{phot}}), \quad p_{\text{son}} = \pm(p_{\text{phot}} - p'_{\text{phot}}),$$

où le signe « plus » concerne l'émission et le signe « moins » l'absorption du phonon. Les quantités non primées dénotent l'énergie et l'impulsion du photon avant émission ou absorption du phonon et celles affectées de « primes » concernent ces mêmes quantités après émission ou absorption du phonon. Après avoir multiplié la deuxième équation par c/n, élevons les deux équations au carré, puis soustrayons membre à membre. En utilisant la relation entre l'énergie et l'impulsion on obtient alors

$$\left(\frac{c^2}{n^2 v^2_{\text{son}}} - 1\right) \mathscr{E}^2_{\text{son}} = 4\mathscr{E}_{\text{phot}} \mathscr{E}'_{\text{phot}} \sin^2 \frac{\vartheta}{2},$$

où ϑ est l'angle entre les directions du photon incident et du photon diffusé. Dans cette dernière équation, on peut négliger l'unité figurant dans la parenthèse du premier membre et remplacer $\mathscr{E}'_{\text{phot}}$ par $\mathscr{E}_{\text{phot}}$ dans le second membre, puisque l'énergie du phonon est extrêmement petite. Après cette simplification et en prenant la racine carrée, on obtient

$$\mathscr{E}_{\text{son}} = \pm 2n \frac{v_{\text{son}}}{c} \mathscr{E}_{\text{phot}} \sin \frac{\vartheta}{2},$$

ou

$$\mathscr{E}_{\text{phot}}-\mathscr{E}'_{\text{phot}}=\pm 2n\,\frac{v_{\text{son}}}{c}\,\mathscr{E}_{\text{phot}}\sin\frac{\vartheta}{2}\,. \tag{57.8}$$

C'est une relation purement classique qui a été obtenue en n'utilisant que les lois de la conservation de l'énergie et de l'impulsion, ainsi que les relations classiques entre l'énergie et l'impulsion pour la lumière et le son. Pour passer de l'énergie à la fréquence, on doit utiliser les relations quantiques $\mathscr{E}_{\text{phot}} = \hbar\omega$ et $\mathscr{E}_{\text{phon}} = \hbar\omega_{\text{phon}}$; on notera que dans ces deux relations, la constante $\hbar$ est la même. Finalement, en adoptant l'interprétation quantique, on retrouve la même formule que fournit la théorie classique:

$$\omega-\omega'=\pm 2n\,\frac{v_{\text{son}}}{c}\,\omega\sin\frac{\vartheta}{2}\,. \tag{57.9}$$

On notera que le résultat final est indépendant de la valeur numérique de la constante de Planck.

§ 58. Bandes d'énergie des corps solides

1. Dans le t. III, § 100, on avait introduit la notion des bandes d'énergie pour les corps cristallins. Comme cette notion se fonde essentiellement sur les conceptions quantiques que l'on ne pouvait utiliser dans le t. III, on reprendra dans ce paragraphe la question des bandes d'énergie. On s'attachera à en exposer l'aspect qualitatif, les principes de base, sans aborder l'étude quantitative qui implique des calculs longs et ardus. D'ailleurs, l'état de la théorie ne permet pas d'effectuer des calculs exacts pour les cristaux réels et on en est réduit à se contenter de modèles simples et de considérations évidentes. Le lecteur est invité à comparer l'exposé de la question donnée dans ce paragraphe avec l'exposé du § 100, t. III.

En principe, on pourrait procéder comme on avait commencé à le faire au § 100 du t. III. Prenons une multitude d'atomes identiques, éloignés les uns des autres à des distances suffisantes pour que les interactions mutuelles des atomes soient tellement faibles qu'on ne puisse déceler leur résultat. Chaque atome se comporte alors comme s'il était seul et sera caractérisé par des niveaux énergétiques qui lui sont propres. Un système de N atomes éloignés les uns des autres possédera ainsi les mêmes niveaux énergétiques, mais chaque niveau sera répété N fois. Commençons maintenant à rapprocher les atomes. Ceux-ci entrent en interaction et chacun des niveaux énergétiques initiaux commence à se déplacer. Finalement, les N atomes formeront un cristal. En suivant l'évolution des différents niveaux énergétiques, on peut déterminer les niveaux énergétiques du cristal qui s'est constitué.

On conçoit que si le nombre N d'atomes est très grand, il est pratiquement impossible de réaliser ce programme. Mais on peut le faire pour deux atomes et utiliser les résultats ainsi obtenus pour se faire une idée qualitative du comportement d'un système formé d'un

grand nombre d'atomes. En outre, pour simplifier encore le problème, on peut remplacer l'atome réel par un « atome unidimensionnel » imaginaire, par exemple par un oscillateur harmonique. C'est ce qu'on a déjà fait au § 52 pour mettre en évidence la nature des forces moléculaires (voir aussi t. III, § 137). Nous allons reproduire les résultats qui y ont été obtenus en les présentant sous une forme adéquate à nos besoins.

2. Tant que les oscillateurs se trouvent à grande distance les uns des autres, l'hamiltonien du système de deux oscillateurs sera représenté par (les notations sont celles du § 52)

$$\hat{H} = \hat{H}_1 + \hat{H}_2, \tag{58.1}$$

où

$$\begin{aligned} \hat{H}_1 &= -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x_1^2} + \frac{\mu}{2}\omega_0^2 x_1^2, \\ \hat{H}_2 &= -\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial x_2^2} + \frac{\mu}{2}\omega_0^2 x_2^2. \end{aligned} \tag{58.2}$$

Les oscillateurs étant indépendants l'un de l'autre, l'énergie de chacun d'eux est quantifiée et vaut $\hbar\omega_0\left(n + \frac{1}{2}\right)$. L'énergie du système est égale à la somme des énergies des deux oscillateurs. Comme les oscillateurs sont identiques, une énergie donnée $\mathcal{E}$ du système peut être représentée de deux façons différentes: $\mathcal{E} = \mathcal{E}_1^{\text{I}} + \mathcal{E}_2^{\text{II}}$ ou $\mathcal{E} = \mathcal{E}_2^{\text{I}} + \mathcal{E}_1^{\text{II}}$. Dans le premier cas, l'oscillateur I possède l'énergie $\mathcal{E}_1$ et l'oscillateur II possède l'énergie $\mathcal{E}_2$. Dans le second cas, l'énergie totale $\mathcal{E}$ est la même, mais l'oscillateur II possède l'énergie $\mathcal{E}_1$ et l'oscillateur I possède l'énergie $\mathcal{E}_2$. Cela signifie que le niveau énergétique $\mathcal{E}$ du système d'oscillateurs est deux fois dégénéré. Lorsqu'on rapproche les oscillateurs, on voit apparaître dans l'hamiltonien du système

$$\hat{H} = \hat{H}_1 + \hat{H}_2 - \lambda x_1 x_2 \tag{58.3}$$

le terme $-\lambda x_1 x_2$ qui tient compte de l'interaction mutuelle des oscillateurs. Pour mettre en évidence le rôle de ce terme, le plus simple est d'introduire les coordonnées normales

$$q_1 = \frac{x_1 + x_2}{\sqrt{2}}, \quad q_2 = \frac{x_1 - x_2}{\sqrt{2}}. \tag{58.4}$$

L'hamiltonien du système devient alors

$$\hat{H} = \left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial q_1^2} + \frac{\mu}{2}\omega_1^2 q_1^2\right) + \left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial q_2^2} + \frac{\mu}{2}\omega_2^2 q_2^2\right), \tag{58.5}$$

où

$$\omega_1^2 = \omega_0^2 - \frac{\lambda}{\mu} \qquad \omega_2^2 = \omega_0^2 + \frac{\lambda}{\mu}. \tag{58.6}$$

La forme de l'hamiltonien (58.5) montre qu'en coordonnées normales q_1 et q_2 le système exécute deux vibrations collectives indépendantes avec les fréquences ω_1 et ω_2. On peut exprimer ce résultat autrement. Formellement, l'introduction des coordonnées normales signifie qu'on décrit le mouvement du système d'oscillateurs à l'aide de deux quasi-particules qui exécutent des oscillations harmoniques de fréquences ω_1 et ω_2. Le mouvement de chacune des quasi-particules décrit non pas le mouvement d'un seul oscillateur, mais le mouvement des deux oscillateurs à la fois. Le mouvement d'ensemble du système résulte de la superposition des mouvements des deux quasi-particules. L'énergie de la première quasi-particule est $\hbar\omega_1\left(n_1 + \frac{1}{2}\right)$ et celle de la seconde $\hbar\omega_2\left(n_2 + \frac{1}{2}\right)$ et aucun de ces niveaux n'est dégénéré. Ainsi, par suite de l'interaction, le niveau énergétique deux fois dégénéré se scinde en deux niveaux de fréquences ω_1 et ω_2. Le nombre total de niveaux ne change pas.

On pourrait répéter sans grande difficulté ces estimations pour un système comportant un nombre quelconque d'oscillateurs harmoniques identiques. Le problème consiste à introduire des coordonnées normales avec lesquelles les énergies cinétique et potentielle du système se laissent ramener simultanément à des sommes de carrés. Cette procédure est bien connue en algèbre linéaire et trouve de nombreuses applications, par exemple en théorie des petites oscillations des systèmes oscillants à grand nombre de degrés de liberté. Le système tout entier se laisse encore formellement ramener à un ensemble de quasi-particules indépendantes (oscillateurs harmoniques) ayant des fréquences propres différentes. Chaque niveau d'énergie initialement dégénéré se scinde en plusieurs niveaux, le nombre total de niveaux restant inchangé lorsqu'on rapproche les oscillateurs éloignés.

3. Maintenant on peut revenir à la question posée au début de ce paragraphe : à quoi peut-on s'attendre si on rapproche des atomes identiques infiniment éloignés, lorsque leur nombre N est infiniment grand ? En position initiale des atomes, ceux-ci n'interagissent pas. De ce fait la multiplicité de dégénérescence de chaque niveau énergétique d'un système formé d'un nombre incalculable d'atomes est démesurée. Lorsqu'on rapproche les atomes, ceux-ci entrent en interaction et la dégénérescence est levée. Par analogie avec le système d'oscillateurs, on peut s'attendre à ce que les niveaux énergétiques d'un système d'atomes se scindent. Finalement, lorsque les atomes auront formé un cristal, chaque niveau énergétique initialement dégénéré éclate en un nombre énorme de niveaux. A la place d'un niveau dégénéré apparaît un nombre énorme de niveaux formant une *bande d'énergie*.

On comprend maintenant comment se forme une structure des bandes d'énergie d'un cristal à partir des niveaux individuels d'ato-

mes isolés. Ces considérations peuvent être étendues, sans modification essentielle, au cas de cristaux formés d'atomes d'espèces différentes (cristaux de NaCl par exemple).

4. L'interaction mutuelle des atomes qui apparaît dès qu'on les rapproche affecte en premier lieu les *couches électroniques externes* dont les fonctions d'onde se recouvrent fortement. L'interaction des atomes affecte beaucoup moins les électrons des couches internes dont les fonctions d'onde ne se recouvrent que fort peu. Il s'ensuit que les niveaux atomiques profonds forment dans le corps solide des bandes très étroites, tandis que les niveaux supérieurs donnent naissance à des bandes d'énergie notablement plus larges. Les bandes les plus larges correspondent aux niveaux des *électrons de valence.* C'est pour cela que par échauffement des corps solides on voit apparaître des spectres optiques continus et non pas des spectres de raies. En effet, l'émission des raies du spectre optique, qui correspondent à de grandes longueurs d'onde, résulte de transitions entre des niveaux d'énergie de la bande extrême, dont les intervalles sont petits et qui forment une suite continue. D'autre part, dans les tubes à rayons X à anticathodes en métaux lourds fonctionnant sous de grandes tensions, on voit apparaître, sur un fond continu, des raies fines du rayonnement caractéristique qui correspondent aux transitions entre des bandes d'énergie internes de faible largeur.

Par conséquent, le spectre énergétique d'un cristal comporte plusieurs bandes dites *bandes d'énergie*, chacune résultant de l'éclatement des niveaux énergétiques des atomes individuels par suite de leurs interactions mutuelles. A ce processus d'éclatement des niveaux participent surtout les électrons de valence, i.e. les électrons périphériques des atomes. Comme le nombre de niveaux résultant de l'éclatement des niveaux énérgétiques des atomes individuels est extrêmement grand, l'énergie du cristal varie de façon pratiquement continue dans les limites de chaque bande. Les bandes d'énergie voisines peuvent être séparées par des interstices (*bandes interdites*) de largeurs finies. L'énergie du cristal ne peut pas prendre les valeurs situées dans la bande interdite. Toutes les autres bandes sont dites *permises*. Il peut arriver aussi que les bandes voisines se touchent ou se chevauchent. La bande extrême formée par les niveaux de plus grande énergie est appelée *bande de conduction*. Les bandes plus profondes qui résultent de l'éclatement des niveaux des électrons de valence sont appelées *bandes de valence*.

5. Rappelons et étendons les considérations qui ont été développées dans le t. III, § 100 sur le mécanisme du passage du courant électrique à travers un cristal. Tous les cristaux sont classés en *métaux*, *diélectriques* et *semiconducteurs*. Afin d'exclure l'influence de l'agitation thermique, on supposera d'abord que la température du cristal est égale au zéro absolu. Selon le principe de Pauli, chaque niveau énergétique ne peut être occupé par plus de deux électrons

ayant des spins antiparallèles. A l'état d'équilibre les niveaux énergétiques les plus bas seront occupés par les électrons, les niveaux supérieurs restant vacants.

Dans les diélectriques, la bande de valence est entièrement occupée par les électrons, tandis que la bande de conduction qui se trouve plus haut dans l'échelle des énergies et qui en est séparée par une bande interdite de largeur notable, ne contient aucun électron. Il s'ensuit que dans des champs électriques faibles (inférieurs au champ de claquage) les diélectriques ne sont pas conducteurs d'électricité. En effet, lorsqu'on applique un champ électrique continu extérieur, à l'hamiltonien du cristal vient s'ajouter un petit terme additionnel représentant l'énergie du cristal dans le champ électrique extérieur. L'équation de Schrödinger du cristal tout entier s'en trouve légèrement modifiée, mais comme le nombre de niveaux énergétiques de la bande reste inchangé, l'application du champ ne provoque qu'un très petit déplacement des niveaux énergétiques, ce qui se traduit par une *polarisation du cristal*. Ce qui importe surtout, c'est qu'après application du champ électrique la bande reste *complètement remplie d'électrons*. L'état de la bande de valence ne peut être modifié, puisque du point de vue quantique toute modification est liée à des transitions d'électrons entre les niveaux d'énergie. Or aucune transition n'est possible, puisque tous les niveaux de la bande de valence sont occupés par des électrons. Le passage du courant électrique implique un déplacement d'électrons lors duquel ils passent continuellement d'un état à un autre. Il s'ensuit que tant que les électrons se trouvent dans une bande de valence entièrement occupée, *ils ne peuvent participer à aucun transport de courant électrique*. Pour que le courant passe, il faut créer dans la bande de valence des niveaux d'énergie vacants, ce qui consiste à faire passer un certain nombre d'électrons de la bande de valence dans la bande de conduction. Un champ électrique inférieur au champ de claquage ne peut le faire (on ne considère pas l'éventualité de passages par effet tunnel de la bande de valence vers la bande de conduction). Dans les champs électriques intenses apparaît le *claquage électrique* des diélectriques.

6. Dans les métaux, la bande de valence est occupée *entièrement* par des électrons et la bande de conduction ne l'est que *partiellement*. Peu importe s'il existe un interstice entre la bande de valence et la bande de conduction. Ces bandes peuvent se toucher ou se recouvrir, mais ce qui compte, c'est l'existence d'états *non occupés par des électrons* dans la bande de conduction. Lorsqu'on applique un champ électrique extérieur $\boldsymbol{E}$, les états quantiques de la bande de conduction en sont légèrement modifiés, mais les états non occupés par des électrons y subsistent. Les électrons peuvent venir occuper ces états vacants. Toute transition liée au déplacement d'un électron dans le sens du champ $\boldsymbol{E}$, i.e. à l'encontre de la force à laquelle il

est soumis, donne lieu à une augmentation de l'énergie du cristal. La transition qui est liée à un déplacement de sens opposé donne lieu à une diminution de l'énergie du cristal. C'est pour cela que les transitions des électrons qui se font à l'encontre du champ $\boldsymbol{E}$ ont une plus grande probabilité de se réaliser que les transitions inverses. Le cristal sera alors parcouru par un courant électrique circulant dans le sens du champ $\boldsymbol{E}$, conformément à la convention admise pour définir le sens du courant électrique.

Si les charges transportées n'étaient pas évacuées des frontières du cristal, elles y créeraient un champ électrique de sens opposé à celui du champ appliqué $\boldsymbol{E}$, et cela entraînerait la cessation du passage de courant. Pour que le courant puisse passer continuellement il faut évacuer les charges afin que le champ appliqué $\boldsymbol{E}$ reste actif. On pourrait aussi créer dans le cristal un champ électrique rotationnel, en le plaçant par exemple dans un champ magnétique alternatif. Dans ce cas aucun équilibre n'est possible et un courant d'induction électrique circulera sans arrêt dans le cristal.

7. Comme dans les diélectriques, dans les semiconducteurs la bande de valence est *entièrement occupée* par les électrons et la bande de conduction est *complètement vide*. Ces deux bandes sont séparées par une bande interdite de largeur finie, mais dans les semiconducteurs elle est beaucoup plus étroite que dans les diélectriques. Au zéro absolu cette dernière circonstance ne joue aucun rôle, puisque les transitions d'électrons de la bande de valence à la bande de conduction sont irréalisables (compte non tenu des transitions par effet tunnel). Mais si la température du cristal est supérieure au zéro absolu, un électron de la bande de valence peut recevoir de l'ion du réseau cristallin une énergie de l'ordre de kT et passer dans la bande de conduction. (Il s'agit bien d'une énergie de l'ordre de kT, parce que la concentration des électrons dans la bande de conduction étant petite, les électrons sont pratiquement régis par la statistique classique de Boltzmann.) Cette transition de l'électron peut être réalisée par d'autres procédés, par exemple, en éclairant le cristal, mais quel que soit le mode de transition des électrons dans la bande de conduction, le cristal acquiert une certaine conductibilité électrique.

Comme dans les métaux, la conductibilité des semiconducteurs est due aux électrons passés dans la bande de conduction. Mais il existe également un autre mécanisme de conduction. Tout électron quittant la bande de valence y laisse un état vacant que l'on appelle *trou*. Un autre électron de la bande de valence peut maintenant venir occuper cet état vacant, laissant derrière lui un nouveau trou, qui à son tour pourra être occupé par un troisième électron, etc. Si le cristal se trouve soumis à l'action d'un champ électrique extérieur $\boldsymbol{E}$, les transitions liées au mouvement des électrons à l'encontre du champ $\boldsymbol{E}$ sont plus probables que les transitions liées aux mouve-

ments de sens opposés. C'est ce qui détermine le passage du courant dans le sens de $\boldsymbol{E}$. Au mouvement de l'électron se superpose le mouvement du trou correspondant, qui s'effectue en sens inverse. Tout se passe comme si le courant était dû non pas au mouvement de d'électrons négatifs, mais à un mouvement de sens inverse de trous chargés positivement.

La présence d'impuretés, même en quantités infimes, détermine un accroissement important de la conductibilité électrique des semiconducteurs. Les impuretés créent dans la bande interdite des niveaux d'énergie supplémentaires; si ces niveaux se trouvent près du bord inférieur de la bande de conduction, les électrons qui les occupent peuvent passer dans la bande de conduction et augmenter la conductibilité du semiconducteur. Les impuretés qui fournissent des électrons à la bande de conduction et font croître la conductibilité du semiconducteur sont appelées des *donneurs*. Si les niveaux supplémentaires se trouvent près du bord supérieur de la bande de valence, ils pourront être occupés par des électrons provenant de la bande de valence, en laissant derrière soi des trous. En conséquence, la conductibilité du semiconducteur sera accrue, mais on dit qu'il s'agit alors d'une *conductibilité par trous*. Les impuretés correspondantes sont appelées des *accepteurs*.

8. La conductibilité des semiconducteurs s'accroît fortement avec la température. Dans les semiconducteurs purs l'intensification de l'agitation termique des électrons accroît la probabilité de leur passage de la bande de valence à la bande de conduction à travers la bande interdite. En présence d'impuretés la probabilité des transitions des électrons des niveaux donneurs dans la bande de conduction, ou des électrons de la bande de valence sur les niveaux accepteurs, augmente également. Tous ces processus font croître la concentration des porteurs de courant, électrons et trous, dans le semiconducteur, et c'est ce qui détermine l'accroissement de sa conductibilité électrique avec la température.

A ce point de vue les métaux se comportent tout autrement. La concentration électronique des métaux est presque indépendante de la température et leur conductibilité dépend surtout de la *longueur de libre parcours* des électrons. Celle-ci est dans une certaine mesure limitée par la présence des impuretés dans le réseau. La partie correspondante de la conductibilité est indépendante de la température. Dans un réseau cristallin parfait et illimité, ne contenant ni impuretés ni défauts de réseau, l'onde plane de de Broglie associée au mouvement de l'électron se propagerait *sans affaiblissement* et *sans diffusion*. La conductibilité électrique d'un tel cristal serait *infiniment grande*. En réalité, le libre parcours de l'électron est limité par les dimensions du cristal (cf. § 57, 4), ainsi que par les fluctuations thermiques et par les défauts du réseau cristallin. Ceux-ci augmentent avec la température et réduisent la longueur de libre parcours de

l'électron. En conséquence, *la conductibilité électrique des métaux purs diminue à mesure que la température augmente.* (Les supraconducteurs seront étudiés plus loin.)

§ 59. Structure de bande et ondes de Bloch

1. On peut également déduire la structure de bande du spectre énergétique d'un cristal à partir de l'équation de Schrödinger pour les états stationnaires. Or il est impossible de résoudre et d'analyser l'équation exacte de Schrödinger d'un cristal, car ce dernier contient un nombre énorme de particules (électrons et noyaux atomiques). Il est donc indispensable de simplifier l'équation de Schrödinger. On admet alors que l'état du cristal peut être décrit de façon approchée par des *fonctions d'onde monoélectroniques* de chacun des électrons au lieu de la fonction d'onde complète dépendant des coordonnées de tous les électrons et de tous les noyaux atomiques. Chacune de ces fonctions d'onde monoélectroniques ne dépend que des coordonnées d'un *seul* électron soumis à l'action du champ de forces créé par les noyaux atomiques et par tous les autres électrons du cristal. Les noyaux étant lourds sont supposés être fixes, tandis que les électrons sont supposés être « étalés » dans tout le volume du cristal. C'est ainsi que l'équation de Schrödinger pour un grand nombre de particules se trouve remplacée par une équation monoélectronique relative à un seul électron.

Il importe de noter que le champ à forces conservatives dans lequel se trouve l'électron n'est pas donné et dépend des états des électrons. On dit que c'est un champ *autocongruent*. Ce résultat s'obtient, d'une part, parce que les fonctions d'onde monoélectroniques de chaque électron sont définies par le champ autocongruent, et, d'autre part, parce que ce dernier dépend lui-même de la forme des fonctions d'onde monoélectroniques des électrons. La justification de la mise en œuvre de la méthode du champ autocongruent réside en ce que la plupart des résultats qu'elle permet d'obtenir sont conformes aux données expérimentales. Cela tient non pas à la forme explicite du champ autocongruent mais à sa *périodicité spatiale* déterminée par la périodicité du réseau cristallin.

Il convient de remarquer encore qu'il suffit d'introduire les fonctions d'onde non pas pour tous les électrons, mais seulement pour les électrons périphériques, i.e. les *électrons de valence* dont la liaison avec les atomes du réseau est relativement faible; on admet que tous les autres électrons sont fortement attachés aux noyaux atomiques. On aboutit ainsi à un modèle cristallin composé d'ions positivement chargés que l'on admet être fixes. Le champ spatialement périodique autocongruent est créé par les ions fixes et par les électrons de valence qui se déplacent librement entre les ions.

2. Prenant pour base ce modèle de cristal, cherchons à déterminer la forme des fonctions d'onde monoélectroniques stationnaires d'un électron du cristal, ainsi que la structure de son spectre énergétique en supposant qu'il n'y a aucun champ de forces extérieur et que le réseau s'étend indéfiniment de tous côtés. Afin d'éviter les complications mathématiques, on considérera un *cristal unidimensionnel* (i.e. une chaîne rectiligne infinie d'atomes identiques avec une distance de séparation a). Dans ce cas l'équation de Schrödinger pour les états stationnaires est de la forme

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+U(x)\psi=\mathscr{E}\psi \tag{59.1}$$

ou

$$\frac{d^2\psi}{dx^2}+\varkappa^2(x)\psi=0 \tag{59.2}$$

avec la notation

$$\varkappa^2=\frac{2m}{\hbar^2}(\mathscr{E}-U). \tag{59.3}$$

C'est une équation différentielle linéaire de second ordre à coefficients périodiquement variables, puisqu'en l'absence de tout champ extérieur on a $\varkappa^2(x+a)=\varkappa^2(x)$ pour tout x. On l'appelle *équation de Hill* (cas particulier de l'*équation de Mathieu*).

Examinons la forme générale des solutions de l'équation de Hill en nous fondant sur la périodicité de la fonction $\varkappa^2(x)$. En raison de cette périodicité

$$\frac{d^2\psi(x+a)}{dx^2}+\varkappa^2(x+a)\psi(x+a)=\frac{d^2\psi(x+a)}{dx^2}+\varkappa^2(x)\psi(x+a).$$

Il en résulte que si la fonction $\psi(x)$ est solution de l'équation de Hill, la fonction $\psi(x+a)$ le sera également. Si $\psi_1(x)$ et $\psi_2(x)$ sont deux solutions quelconques, linéairement indépendantes de l'équation de Hill, sa solution générale sera de la forme

$$\psi(x)=c_1\psi_1(x)+c_2\psi_2(x),$$

où c_1 et c_2 sont des constantes arbitraires.

Démontrons que parmi les solutions de l'équation de Hill il existe une solution $\Phi(x)$ telle que $\Phi(x+a)=\lambda\Phi(x)$ pour tous les x, λ étant une constante. Si cette solution existe bien, on peut évidemment la mettre sous la forme

$$\Phi(x)=c_1\psi_1(x)+c_2\psi_2(x).$$

Pour simplifier les calculs, prenons des solutions linéairement indépendantes $\psi_1(x)$ et $\psi_2(x)$ telles qu'elles vérifient les conditions suivantes:

$$\psi_1(0)=1,\qquad \psi_1'(0)=0;$$
$$\psi_2(0)=0,\qquad \psi_2'(0)=1.$$

(On dit dans ce cas que les fonctions $\psi_1(x)$ et $\psi_2(x)$ forment un *système de solutions fondamental.*) Dans ce cas pour $x = 0$, les conditions $\Phi(x + a) = \lambda\Phi(x)$ et $\Phi'(x + a) = \lambda\Phi'(x)$ s'écrivent

$$c_1\psi_1(a) + c_2\psi_2(a) = \lambda c_1,$$
$$c_1\psi_1'(a) + c_2\psi_2'(a) = \lambda c_2$$

ou encore

$$[\psi_1(a) - \lambda]\, c_1 + \psi_2(a)\, c_2 = 0,$$
$$\psi_1'(a)\, c_1 + [\psi_2'(a) - \lambda]\, c_2 = 0.$$

Pour que ces équations linéaires et homogènes (par rapport à c_1 et c_2) soient compatibles il faut et il suffit que soit vérifiée la condition

$$\begin{vmatrix} \psi_1(a)-\lambda & \psi_2(a) \\ \psi_1'(a) & \psi_2'(a)-\lambda \end{vmatrix} = 0 \qquad (59.4)$$

ou

$$\lambda^2 - [\psi_1(a) + \psi_2'(a)]\,\lambda + [\psi_1(a)\,\psi_2'(a) - \psi_1'(a)\,\psi_2(a)] = 0.$$

En différentiant, compte tenu de l'équation (59.2), il est facile de s'assurer que

$$\frac{d}{dx}[\psi_1(x)\,\psi_2'(x) - \psi_1'(x)\,\psi_2(x)] = 0.$$

Il s'ensuit que la fonction entre crochets est constante. Pour trouver sa valeur on posera $x = 0$, et la fonction devient alors égale à 1. Cette fonction conservera cette valeur pour $x = a$. Introduisons la notation

$$\psi_1(a) + \psi_2'(a) = 2L.$$

Bien entendu, la quantité L est constante, puisqu'on utilise la solution fondamentale de l'équation de Hill (59.2). On l'appelle *constante de Liapounov* (1857-1918). Sa valeur est déterminée par la fonction $\varkappa$, donc en fait par le paramètre $\mathscr{E}$. Maintenant on déterminera λ par l'équation du second degré

$$\lambda^2 - 2L\lambda + 1 = 0. \qquad (59.5)$$

On en tire deux valeurs de λ:

$$\lambda_{1,\,2} = L \pm \sqrt{L^2 - 1}. \qquad (59.6)$$

On obtient ainsi non pas une mais *deux* solutions de l'équation (59.2) : $\Phi_1(x)$ et $\Phi_2(x)$ possédant la propriété requise. (On s'abstiendra d'examiner le cas où les racines sont égales puisqu'on peut l'analyser en effectuant le passage à la limite $\lambda_1 \to \lambda_2$.) Les solutions $\Phi_1(x)$ et $\Phi_2(x)$ étant linéairement indépendantes, il est commode de les utiliser pour représenter la solution générale sous la forme

$$\psi = c_1\Phi_1(x) + c_2\Phi_2(x). \qquad (59.7)$$

On conçoit qu'il est impossible de trouver les solutions $\Phi_1(x)$ et $\Phi_2(x)$ tant qu'on ne connaît pas la fonction $\varkappa^2(x)$ (en fin de compte tant qu'on n'aura pas explicité le champ autocongruent U), et que le paramètre $\mathcal{E}$ n'aura pas été fixé. Néanmoins, on peut opérer avec ces fonctions pour trouver les propriétés générales des solutions de l'équation de Hill (59.2). Or, les propriétés intéressantes de ces solutions dépendent essentiellement de la valeur de la constante de Liapounov L.

3. Supposons d'abord qu'en module, la constante L soit supérieure à l'unité: $|L| > 1$. Dans ce cas les deux racines λ_1 et λ_2 sont réelles, l'une étant, en module, plus grande et l'autre, plus petite que l'unité puisque $\lambda_1\lambda_2 = 1$. Il résulte de l'équation $\Phi(x + a) = \lambda\Phi(x)$ que pour $x \to +\infty$ l'une des fonctions $\Phi_1(x)$ ou $\Phi_2(x)$ croît indéfiniment, tandis que pour $x \to -\infty$ c'est l'autre fonction qui croît indéfiniment. Il s'ensuit que pour $|L| > 1$ aucune des solutions $\Phi_1(x)$ ou $\Phi_2(x)$, et par suite aucune de leurs combinaisons linéaires à coefficients constants (59.7) ne peut être la fonction d'onde d'un électron dans le cristal. Cela signifie qu'il ne peut exister dans le cristal d'états d'énergie $\mathcal{E}$ pour lesquels $|L| > 1$. Ces valeurs de l'énergie correspondent aux *bandes interdites*.

Lorsque $|L| < 1$ on peut poser $L = \cos ka$, k étant une constante. On a alors

$$\lambda_{1,2} = \cos ka \pm i \sin ka = e^{\pm ika}.$$

Il s'ensuit que $\Phi_{1,2}(x + a) = e^{\pm ika}\,\Phi_{1,2}(x)$, ce qui signifie que lorsque x varie de a, le module de la fonction $\Phi(x)$ reste constant, et seule sa phase varie. Les deux fonctions $\Phi_1(x)$ et $\Phi_2(x)$ ainsi que toutes leurs combinaisons linéaires à coefficients constants peuvent être des solutions de l'équation de Schrödinger (59.1). Par conséquent ne sont possibles que les valeurs de l'énergie $\mathcal{E}$ pour lesquelles le module de la constante de Liapounov est inférieur à l'unité. Ces valeurs de $\mathcal{E}$ forment des *bandes d'énergie permises d'un cristal.*

Ces développements nous ont conduit à la *structure de bande du spectre énergétique d'un électron dans un cristal*, qui a été déjà décrite dans le paragraphe précédent. Dans les limites d'une bande donnée, l'énergie d'un électron varie de façon continue. Ce résultat est lié à l'hypothèse que *la chaîne est infinie*. Dans le cas d'une chaîne limitée, à ses extrémités devraient être satisfaites certaines conditions aux limites, qui auraient pour conséquence le remplacement du spectre continu de chaque bande par une suite de niveaux plus ou moins rapprochés. Le problème n° 1 donné à la fin de ce paragraphe aidera à fixer les idées sur cette question.

4. Les fonctions d'onde $\Phi_1(x)$ et $\Phi_2(x)$ peuvent être présentées sous la forme

$$\Phi_{1,2}(x) = e^{\pm ikx} P_{1,2}(x).$$

De toute évidence on a d'une part

$$\Phi_{1,2}(x+a) = e^{\pm ik(x+a)} P_{1,2}(x+a)$$

et d'autre part

$$\Phi_{1,2}(x+a) = e^{\pm ika}\Phi_{1,2}(x) = e^{\pm ik(x+a)} P_{1,2}(x).$$

Par conséquent,

$$P_{1,2}(x+a) = P_{1,2}(x), \tag{59.8}$$

autrement dit les deux fonctions $P_1(x)$ et $P_2(x)$ sont périodiques de période a.

Si on tient compte encore du facteur temporel $e^{-i\omega t}$ et s'il n'y a aucun champ de forces extérieur, les fonctions d'onde complètes du cristal peuvent être de la forme

$$\begin{aligned} \Psi_1(x, t) &= P_1(x)\, e^{-i(\omega t - kx)}, \\ \Psi_2(x, t) &= P_2(x)\, e^{-i(\omega t + kx)}. \end{aligned} \tag{59.9}$$

Ces ondes décrivent le « mouvement libre » de l'électron dans le cristal, lorsqu'il n'est soumis qu'aux forces de ses interactions avec les ions du réseau cristallin et les autres électrons, aucun champ de forces extérieur n'agissant sur lui. Ces ondes sont appelées *ondes de Bloch* (né en 1905). A la différence des ondes planes de de Broglie qui se propagent dans l'espace libre, les quantités $P_1(x)$ et $P_2(x)$ figurant dans les expressions des ondes de Bloch ne sont pas constantes mais *spatialement modulées*, i.e. varient périodiquement le long de la chaîne avec la période a. Du fait de cette modulation spatiale, dans le cas d'une propagation libre des ondes Ψ dans un cristal, la quantité $\hbar k$ est appelée *quasi-impulsion de l'électron*, tandis que dans le cas du mouvement de l'électron dans l'espace libre cette quantité est son impulsion. Dans les cristaux tridimensionnels, l'onde plane de Bloch est de la même forme que (59.9), mais $P(x)$ doit être remplacée par la fonction $P(\mathbf{r})$ ayant la même périodicité spatiale que le réseau cristallin, et le nombre d'onde k par le vecteur d'onde $\mathbf{k}$, auquel correspond la *quasi-impulsion* $\hbar\mathbf{k}$.

5. La fonction d'onde de Bloch se trouve déterminée (à un facteur constant près) par le choix de la constante de Liapounov, et donc de la quasi-impulsion $\hbar k$. Simultanément se trouve déterminée l'énergie $\mathscr{E}$ de l'électron figurant dans l'équation stationnaire de Schrödinger sous forme d'un paramètre constant. Il s'ensuit que, dans les limites de chacune des bandes, les valeurs permises de l'énergie de l'électron peuvent être données en fonction de la quasi-impulsion :

$$\mathscr{E} = \mathscr{E}(p). \tag{59.10}$$

On en déduit

$$\omega = \omega(k), \tag{59.11}$$

puisque $\mathscr{E} = \hbar\omega$. Ces deux relations constituent ce qu'il est convenu d'appeler les *lois de dispersion* des ondes électroniques ou de l'électron dans les cristaux. Les lois de dispersion caractérisent l'interaction de l'électron considéré avec les ions du réseau cristallin et avec tous les autres électrons (dont la charge est « étalée » dans l'espace).

Pour une onde de Bloch plane

$$\Psi(x, t) = P(x)\, e^{i(kx-\omega t)} \tag{59.12}$$

le nombre d'onde k n'est pas défini univoquement. La raison en est cependant différente de celle qui concerne les ondes acoustiques se propageant dans le réseau cristallin, puisque dans ce dernier cas la non-univocité de k est liée au *caractère discret des valeurs que peut prendre la coordonnée* x (cf. § 56). Dans le cas de l'onde de Bloch la coordonnée x varie de façon *continue*, mais l'« amplitude » de l'onde $P(x)$ est une fonction de x. On peut toujours modifier la constante k et la fonction $P(x)$ de façon que la fonction d'onde (59.12) reste inchangée. Il suffit pour cela de réaliser les substitutions suivantes:

$$k = k' + \frac{2\pi}{a} p, \quad P'(x) = P(x)\, e^{i2\pi px/a},$$

où p est un nombre entier quelconque. L'équation (59.12) se transforme alors en

$$\Psi(x, t) = P'(x)\, e^{i(k'x-\omega t)},$$

i.e. une onde d'amplitude $P'(x)$ périodiquement variable et de nombre d'onde k'. Cette transformation n'affecte en rien le contenu physique de l'équation.

Lorsqu'on étudie le mouvement de l'électron dans une bande quelconque, on peut mettre à profit le caractère non univoque de k et choisir le nombre d'onde k (ou le vecteur d'onde $\mathbf{k}$) de telle manière que son module prenne *la plus petite valeur* possible. Un cas particulièrement important est celui où $k \ll 2\pi/a$ et l'« amplitude » $P(x)$ contient un terme constant notablement plus grand que tous les autres termes rapidement variables dans l'espace. (On verra par la suite que cette situation correspond au cas où l'électron se meut à proximité d'un des bords de la bande considérée.) Dans ces conditions la fonction Ψ représente une onde d'amplitude constante faiblement modulée avec la période du réseau. Chaque fois qu'on étudie le mouvement moyen de l'électron on peut s'abstraire de cette modulation et utiliser une onde électronique d'amplitude constante, que l'on déduit de $P(x)$ en calculant sa moyenne spatiale. La longueur $\Lambda \equiv 2\pi/k$ d'une telle onde moyenne est supposée être très grande devant la période a du réseau. On peut construire un paquet d'ondes moyennes dans un domaine étroit Δk. La vitesse du *mouvement moyen de l'électron* sera alors égale à la *vitesse de groupe*

$$v = \frac{d\mathscr{E}}{dp}. \tag{59.13}$$

6. Considérons une bande permise quelconque et examinons les relations que l'on peut établir à proximité de ses bords. Sur le bord inférieur de la bande l'énergie $\mathcal{E}$ est minimale et sur le bord supérieur elle est maximale. Par conséquent, la dérivée $d\mathcal{E}/dp$ et la vitesse moyenne de l'électron sur les deux bords de la bande deviennent égales à zéro. C'est pourquoi près des bords de chaque bande le développement en série de $\mathcal{E}$ suivant les puissances de $p = \hbar k$ ne peut commencer que par le terme quadratique. Si on compte l'énergie à partir du bord d'une bande et si on pose que les quasi-impulsions sont nulles aux points de maximum et de minimum, on obtient les résultats suivants:

$$\mathcal{E} = p^2/2m_{\text{eff}}, \quad (59.14)$$

où

$$\frac{1}{m_{\text{eff}}} = \left(\frac{d^2\mathcal{E}}{dp^2}\right)_{p=0}. \quad (59.15)$$

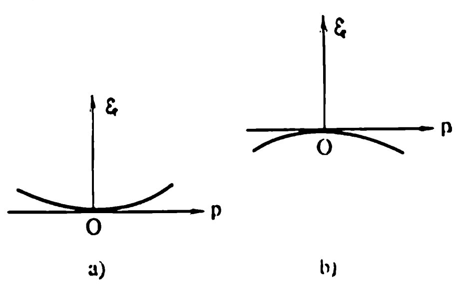


Fig. 105

La quantité m_{eff} est la *masse effective de l'électron*, qui est généralement différente de la masse vraie m_e de l'électron. Sur le bord inférieur des bandes (minimum de $\mathcal{E}$) la masse effective m_{eff} est *positive*, et sur le bord supérieur (maximum de $\mathcal{E}$) elle est *négative* (fig. 105). Dans un réseau cristallin tridimensionnel on remplace $1/m_{\text{eff}}$ par le tenseur $\partial^2\mathcal{E}/\partial p_i \partial p_j$, mais l'essence des phénomènes liés à la notion de masse effective se laisse parfaitement mettre en évidence à l'aide du modèle unidimensionnel où m_{eff} est un scalaire.

Près du bord d'une bande, la relation entre l'énergie de l'électron et sa quasi-impulsion se présente formellement sous la même forme que la relation entre l'énergie et l'impulsion d'une particule libre. C'est ce qui permet de désigner la quantité m_{eff} sous le nom de masse effective et de dénommer la quasi-impulsion p par le terme impulsion tout court. La spécificité du mouvement moyen de l'électron dans un réseau cristallin périodique à proximité du bord d'une bande réside en ce que la masse effective de l'électron peut être positive ou négative. En conséquence, *près du bord inférieur d'une bande les vecteurs* $\mathbf{v}$ *et* $\mathbf{p}$ *sont dirigés dans le même sens et près du bord supérieur ils sont dirigés en sens opposés*, conformément à la formule (59.13) et à la figure 105.

7. On a supposé jusqu'ici qu'il n'y avait aucun champ de forces extérieur. Appliquons maintenant au cristal un champ électrique continu faible. En principe, l'application de ce champ doit déplacer légèrement les niveaux énergétiques du cristal, mais comme ces niveaux se succèdent de manière pratiquement continue, leur disposition à l'intérieur de la bande reste inchangée. Seuls les bords de la bande peuvent subir un léger déplacement, mais si le champ appli-

qué est faible, cet effet est imperceptible et ne se manifeste d'aucune façon. Cependant l'application d'un champ électrique peut modifier l'occupation des niveaux énergétiques par les électrons, à condition, toutefois, qu'il y ait des niveaux énergétiques *occupés partiellement*. Selon le principe de Pauli un niveau ne peut être occupé, au plus, que par deux électrons dont les spins sont antiparallèles. Si la bande est entièrement occupée, les électrons peuvent échanger leurs états, mais comme ils sont indiscernables, cela n'entraîne aucune conséquence physique. Dans ce cas aucun déplacement d'électrons, lié à leurs transitions d'un état à un autre, ne peut se produire, et par suite l'application d'un champ électrique ne peut donner lieu au passage d'un courant électrique.

A l'état d'équilibre et en l'absence de tout champ extérieur, dans une bande partiellement occupée par les électrons tous les niveaux d'énergie *inférieurs* seront occupés chacun par deux électrons de spins opposés. Si on applique un champ électrique, les électrons peuvent passer des niveaux occupés sur des niveaux vacants, ce qui se traduira par le *passage d'un courant électrique*. La vitesse du mouvement moyen des électrons est alors donnée par l'expression (59.13). L'influence qu'exerce le milieu cristallin sur le mouvement des électrons est déjà prise en compte par la formule de dispersion (59.10). Or un électron soumis à l'action d'un champ électrique E subit l'action d'une force extérieure $F = -eE$. Sous l'action de cette force, la variation que subit l'énergie de l'électron au cours du temps dt est égale à $d\mathcal{E} = Fv\,dt$. En vertu de (59.10) $d\mathcal{E} = (d\mathcal{E}/dp)\,dp = v\,dp$. En identifiant les deux expressions, il vient $dp = F\,dt$, i.e.

$$\frac{dp}{dt} = F. \tag{59.16}$$

Dans le cas tridimensionnel on arrive à la même formule, à cette différence près que les scalaires p et F doivent être remplacés par les vecteurs $\boldsymbol{p}$ et $\boldsymbol{F}$. On obtient alors une formule correspondant à la formule classique. En dérivant (59.13) par rapport au temps, on trouve $\dot{v} = (d^2\mathcal{E}/dp^2)\,(dp/dt)$ et en utilisant les formules (52.15) et (52.16), il vient

$$m_{eff}\dot{v} = F. \tag{59.17}$$

8. Considérons le cas particulier où les électrons occupent une bande *presque entièrement*. Le passage du courant électrique à travers le cristal est alors lié à l'existence d'états vacants se trouvant à proximité du bord supérieur de la bande (fig. 105, *b*) où la masse effective de l'électron est négative. Selon (59.17) l'accélération $\dot{v}$ de l'électron est dirigée à l'encontre de la force appliquée $F = eE$, donc dans le sens du champ E ($e < 0$). L'électron se comporte comme une *particule chargée négativement mais ayant une masse m_{eff} négative.*

Or une particule imaginaire ayant une masse et une charge positives se comporterait exactement de la même façon. Il importe peu quels signes on attribuera à la masse et à la charge de cette particule imaginaire, puisque la seule chose qui compte est que la masse et la charge soient de même signe. Un électron de charge et de masse positives se comporte exactement comme le « trou » qui a été défini plus haut. Comme le nombre d'électrons participant au transport du courant électrique est exactement égal au nombre de places vacantes (trous) existant dans la bande, *on peut admettre tout formellement que les porteurs de courant sont les trous.* On comprend maintenant pourquoi le coefficient de Holl, par exemple, peut être non seulement négatif mais aussi positif (cf. t. III, § 98).

9. Notons, pour conclure, que dans un réseau cristallin parfait à ions fixes, l'onde plane de Bloch se propagerait *sans subir d'affaiblissement*, et la résistance électrique du cristal serait alors nulle. L'existence de fluctuations thermiques, de défauts et d'impuretés provoque la diffusion des ondes électroniques, limite donc la longueur de libre parcours des électrons et fait apparaître une résistance électrique.

PROBLÈMES

1. Etudier le comportement d'un réseau cristallin unidimensionnel rectiligne infiniment long en admettant que son énergie potentielle $U(x)$ se laisse

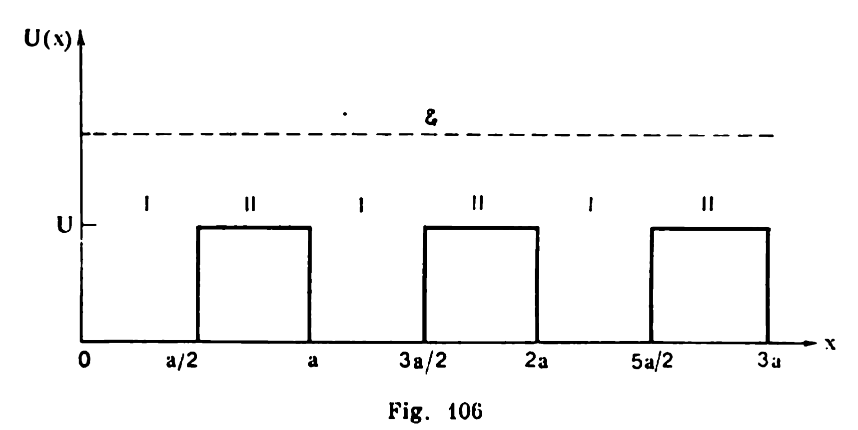


Fig. 106

représenter par la fonction en escalier (fig. 106). Déterminer les bandes permises et interdites de cette chaîne d'atomes en se fixant les valeurs de a et de U correspondant aux dimensions atomiques.

Solution. Posons que l'énergie potentielle sur les segments I est nulle et qu'elle a une valeur constante U sur les segments II. L'équation stationnaire de Schrödinger pour une énergie constante $\mathscr{E}$ s'écrit pour les segments I

$$\frac{d^2\psi}{dx^2}+\varkappa_1^2\psi=0 \qquad (59.18)$$

et pour les segments II elle s'écrit

$$\frac{d^2\psi}{dx^2}+\varkappa_2^2\psi=0, \qquad (59.19)$$

où $\varkappa_1$ et $\varkappa_2$ sont des constantes définies par

$$\varkappa_1^2=\frac{2m}{\hbar^2}\mathscr{E}, \quad \varkappa_2^2=\frac{2m}{\hbar^2}(\mathscr{E}-U).$$

Supposons d'abord que $\mathscr{E} > 0$ et $\mathscr{E} - U > 0$. Dans ce cas $\varkappa_1$ et $\varkappa_2$ seront réels. Sans restreindre la généralité, on peut les considérer comme positifs. Dans l'intervalle $(0, a/2)$ le système des solutions fondamentales est représenté par les fonctions

$$\psi_1=\cos\varkappa_1 x, \quad \psi_2=\frac{1}{\varkappa_1}\sin\varkappa_1 x.$$

Déterminons ces fonctions dans l'intervalle $(a/2, a)$. Dans cet intervalle représentons la première fonction sous la forme

$$\psi_1 = A\cos\varkappa_2(x-a/2)+B\sin\varkappa_2(x-a/2).$$

On détermine les coefficients inconnus A et B à partir des conditions de continuité de la fonction $\psi_1(x)$ et de sa dérivée $\psi_1'(x)$ à la frontière $x = a/2$ de l'intervalle. O trouve ainsi que dans l'intervalle $(a/2, a)$

$$\psi_1(x)=\cos\frac{\varkappa_1 a}{2}\cos\varkappa_2(x-a/2)-\frac{\varkappa_1}{\varkappa_2}\sin\frac{\varkappa a}{2}\sin\varkappa_2(x-a/2).$$

On trouve de même dans le même intervalle

$$\psi_2(x)=\frac{1}{\varkappa_1}\sin\frac{\varkappa_1 a}{2}\cos\varkappa_2(x-a/2)+\frac{1}{\varkappa_2}\cos\frac{\varkappa_1 a}{2}\sin\varkappa_2(x-a/2).$$

La constante de Liapounov est

$$L=\frac{1}{2}[\psi_1(a)+\psi_2'(a)]=$$

$$=\cos\frac{\varkappa_1 a}{2}\cos\frac{\varkappa_2 a}{2}-\frac{1}{2}\left(\frac{\varkappa_1}{\varkappa_2}+\frac{\varkappa_2}{\varkappa_1}\right)\sin\frac{\varkappa_1 a}{2}\sin\frac{\varkappa_2 a}{2}. \qquad (59.20)$$

Dans le cas où $\mathscr{E} > 0$, mais $\mathscr{E} - U < 0$ on modifiera les notations en remplaçant $\varkappa_2$ par la quantité imaginaire $i\varkappa_2$, ce qui revient à poser $\varkappa_2^2 = 2m(U-\mathscr{E})/\hbar^2$. On devra remplacer les fonctions trigonométriques de l'argument imaginaire par des fonctions hyperboliques. La formule (59.20) devient alors

$$L=\cos\frac{\varkappa_1 a}{2}\operatorname{ch}\frac{\varkappa_2 a}{2}-\frac{1}{2}\left(\frac{\varkappa_1}{\varkappa_2}-\frac{\varkappa_2}{\varkappa_1}\right)\sin\frac{\varkappa_1 a}{2}\operatorname{sh}\frac{\varkappa_2 a}{2}. \qquad (59.21)$$

Enfin, dans le cas où $\mathscr{E} < 0$ et $\mathscr{E} - U < 0$, on devra réaliser encore la substitution $\varkappa_1 \to i\varkappa_2$ (i.e. poser $\varkappa_1^2 = -2m\mathscr{E}/\hbar^2$). On obtient alors

$$L=\operatorname{ch}\frac{\varkappa_1 a}{2}\operatorname{ch}\frac{\varkappa_2 a}{2}+\frac{1}{2}\left(\frac{\varkappa_1}{\varkappa_2}+\frac{\varkappa_2}{\varkappa_1}\right)\operatorname{sh}\frac{\varkappa_1 a}{2}\operatorname{sh}\frac{\varkappa_2 a}{2}. \qquad (59.22)$$

Les formules (59.20) à (59.22) ont une forme assez compliquée, aussi est-il recommandé de procéder à leur analyse par un procédé graphique en considérant des cas particuliers. Donnons un exemple numérique en posant approximativement que $a = 2\cdot10^{-8}$ cm, $U = 5$ eV. La courbe correspondante pour la constante $L = L(\mathscr{E})$ est représentée sur la figure 107; sur les parties hachurées la quantité $|L|$ est inférieure à l'unité. Pour notre modèle ces parties hachurées

correspondent aux bandes permises, tandis que les parties non hachurées où $|L| > 1$ correspondent aux bandes interdites.

2. Entre les pôles d'un accumulateur est branché un cristal, dont la bande de conduction est presque entièrement remplie d'électrons. La masse effective

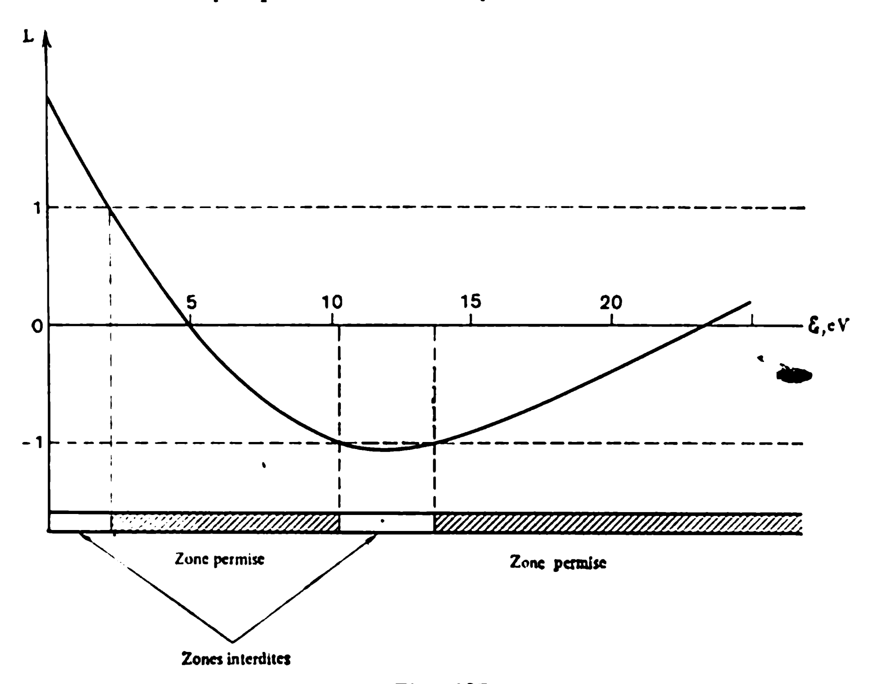


Fig. 107

Fig. 108

des électrons étant alors négative, ils se déplacent à travers le cristal dans le sens du champ électrique. La vitesse moyenne de l'électron est dirigée dans le même sens, puisqu'elle était nulle avant l'application du champ. L'électron se déplace donc à travers le cristal en allant de l'anode vers la cathode. A l'intérieur de l'accumulateur les électrons se déplacent de la cathode vers l'anode, et suivant les conventions en vigueur le courant y circule en sens opposé — de l'anode vers la cathode. Or un courant de ce sens provoquerait la charge de l'accumulateur, ce qui est en contradiction avec la loi de la conservation de l'énergie. Le lecteur est invité à résoudre ce paradoxe.

Solution. Dans l'énoncé ci-dessus on n'a pas tenu compte de la redistribution des électrons parmi les différents états disponibles résultant de l'application du champ électrique. La figure 108 représente une partie de la courbe de dispersion du cristal. En champ nul la droite horizontale en pointillé AB intercepte sur cette courbe la partie supérieure non remplie par les électrons. Dans les parties occupées il y a autant d'électrons à se déplacer à droite qu'il y en a à se déplacer à gauche. En conséquence, il ne circule aucun courant à travers le cristal. Lorsqu'on applique un champ électrique, l'électron se trouve soumis à l'action de la force $F = eE$ dirigée à l'encontre du champ E puisque la charge de l'électron est négative. Supposons, pour fixer les idées, que le champ est orienté de droite à gauche, ce qui implique que la force F est orientée de gauche à droite. Selon (59.16) les impulsions des électrons allant de gauche à droite doivent augmenter en module et celles des électrons allant dans le sens opposé doivent diminuer. Cela signifie que l'extrémité B de la partie de la courbe de dispersion dépourvue d'électrons doit s'abaisser en B' et l'extrémité A doit se relever en A'. La vitesse des électrons est orientée dans le sens opposé à celui de p ; sur la figure, dans la partie de gauche elle est dirigée de gauche à droite et dans le sens opposé dans la partie de droite. Or dans la partie de gauche il y a plus d'électrons que dans la partie de droite, ce qui implique que le flux d'électrons sera dirigé dans le cristal de gauche à droite, donc dans le sens de la force F (à l'encontre du champ E). Ce courant devra décharger l'accumulateur au lieu de le charger.

Le même raisonnement s'applique au cas où une petite partie inférieure de la bande de conduction est occupée par des électrons.

§ 60. Superfluidité. Faits expérimentaux

1. La cristallisation (ou solidification) des substances se produit sous l'action des forces d'attraction s'exerçant entre les atomes (ou les molécules), mais le mouvement désordonné de ces particules s'y oppose. Au point de vue classique, ce mouvement est de nature purement thermique et cesse complètement au zéro absolu des températures. En se plaçant à ce point de vue, on doit conclure qu'à l'approche du zéro absolu n'importe quelle substance devrait se solidifier. Au point de vue de la mécanique quantique, les choses se passeraient différemment. Outre les mouvements d'agitation thermique les atomes exécutent encore des *vibrations de zéro*, qui ne disparaissent pas même au zéro absolu. La fréquence ω de ces vibrations (et donc l'énergie au point zéro $\hbar\omega/2$) est d'autant plus grande que la masse de l'atome est petite ($\omega \sim 1/\sqrt{m}$). Si les forces d'attraction mutuelle des atomes sont trop faibles pour s'opposer aux vibrations de zéro, la substance ne se solidifiera pas et restera à l'état liquide même au zéro absolu, si seulement la pression n'est pas très grande. C'est ainsi que se comporte l'hélium et il y a deux raisons à cela. Premièrement, de tous les gaz rares l'hélium est celui dont les atomes ont la plus petite masse. Deuxièmement, les gaz rares possèdent des enveloppes électroniques externes complètes ayant une symétrie sphérique. La couche électronique K, l'unique que possède l'atome d'hélium et qui ne comporte que deux électrons, jouit de ces propriétés. De ce fait les forces d'attraction mutuelle des atomes He sont relative-

ment faibles et la densité de l'hélium liquide est petite (à 2,20 K et à la pression de la vapeur saturante la densité n'est que de 0,146 g/cm³). C'est pour cela qu'à la pression normale l'hélium reste à l'état liquide au zéro absolu. Comme c'est une manifestation des *propriétés quantiques* de l'hélium liquide, on l'appelle *liquide quantique*. L'hélium est la seule substance connue jouissant de la propriété de rester à l'état liquide au zéro absolu. On peut cependant réaliser la solidification de l'hélium en augmentant la pression extérieure jusqu'à 25,3 atm et au-delà. L'augmentation de la pression est pour ainsi dire équivalente à un accroissement des forces d'attraction mutuelle des atomes d'hélium.

Il existe deux isotopes stables d'hélium : ${}_2^4\mathrm{He}$ et ${}_2^3\mathrm{He}$. L'air contient 0,0005 % en volume d'hélium. L'hélium naturel se compose presque en totalité de ${}_2^4\mathrm{He}$, la teneur en ${}_2^3\mathrm{He}$ étant négligeable. Par exemple, dans l'hélium extrait de l'air à côté de ${}_2^4\mathrm{He}$ on ne trouve que 10^{-7} % de ${}_2^3\mathrm{He}$. On ne put étudier sérieusement les propriétés de ${}_2^3\mathrm{He}$ qu'après qu'on apprit à l'obtenir en quantités macroscopiques par des procédés artificiels. On l'obtient dans la réaction des désintégrations β du tritium :

$$ {}_1^3\mathrm{H} \rightarrow {}_2^3\mathrm{He} + e^-. $$

Le tritium ${}_1^3\mathrm{H}$ est obtenu à son tour par une autre réaction nucléaire :

$$ {}_3^6\mathrm{Li} + {}_0^1\mathrm{n} \rightarrow {}_1^3\mathrm{H} + {}_2^4\mathrm{He} $$

qui résulte du bombardement des noyaux ${}_3^6\mathrm{Li}$ avec des neutrons. On connaît d'autres réactions nucléaires libérant l'isotope ${}_2^3\mathrm{He}$.

2. La liquéfaction de l'hélium naturel fut réalisée pour la première fois en 1908 par Kamerlingh Onnes (1853-1926) dans le laboratoire de Leyde en Hollande. C'est à la suite de cette réalisation qu'il devint possible d'étudier les propriétés des substances à des températures proches du zéro absolu. La température critique de ${}_2^4\mathrm{He}$ est $T_{cr} = 5{,}25$ K et ce n'est qu'au-dessous de cette température que l'hélium peut exister à l'état liquide. Sous la pression de 1 atm ${}_2^4\mathrm{He}$ bout à 4,21 K. La température critique de l'isotope léger ${}_2^3\mathrm{He}$ est encore plus basse ($T_{cr} = 3{,}35$ K) ; ${}_2^3\mathrm{He}$ bout à 3,195 K sous une pression de 1 atm. L'hélium naturel ne fut solidifié qu'en 1926, par Keesom (1876-1956), le successeur de Kamerlingh Onnes à la direction du laboratoire de Leyde. Au zéro absolu la solidification de l'hélium n'est possible que sous une pression de 25,3 atm, et à 60 K la pression extérieure doit être de l'ordre de 10 000 atm. L'hélium solide est un corps absolument transparent et incolore. Dans ce qui suit, chaque fois qu'il sera question des propriétés de l'hélium liquide il s'agira de l'isotope lourd ${}_2^4\mathrm{He}$. Ce n'est qu'au paragraphe suivant qu'on donnera quelques renseignements sur les propriétés de l'isotope léger ${}_2^3\mathrm{He}$. Le diagramme d'état de ${}_2^4\mathrm{He}$ est représenté sur la figure 109.

Lorsqu'on abaisse la température l'hélium liquide subit une transformation de phase de second ordre réversible, i.e. une transformation qui ne s'accompagne d'aucun dégagement ou absorption de chaleur, le volume massique du liquide restant constant (cf. t. II, § 120). La température T_λ de cette transformation réversible diminue linéairement avec l'accroissement de la pression de la valeur maximale de 2,17 K sous la pression de la vapeur saturante de l'hélium (37,80 mm Hg) jusqu'à la valeur minimale de 1,77 K sous une

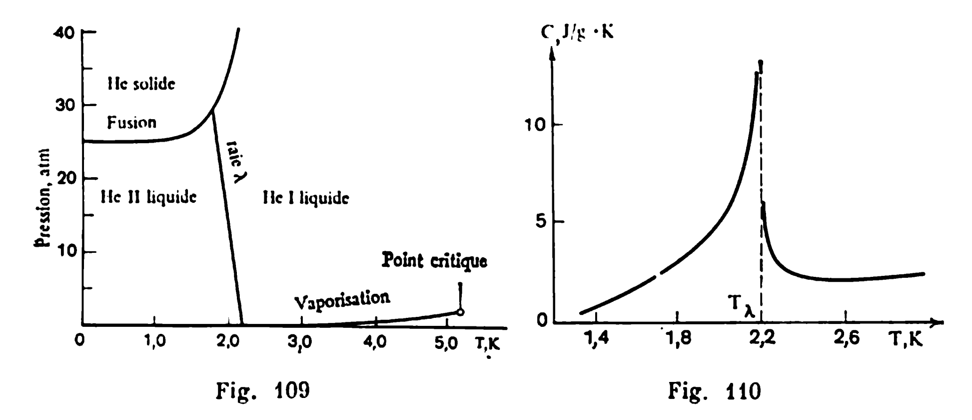


Fig. 109 Fig. 110

pression de 30 atm. Sur le diagramme d'état (T, p) les points figuratifs de cette transformation de phase se situent sur une droite appelée *ligne* λ. Le liquide ayant une température supérieure à T_λ (sous une pression donnée) est appelé *hélium I*, et le liquide dont la température est inférieure à T_λ est appelé *hélium II*. La transition de phase se manifeste premièrement par un *saut de la capacité calorifique* à la température T_λ. De part et d'autre de cette température la capacité calorifique de l'hélium tend vers l'infini selon la loi logarithmique $C = \text{const} \ln \frac{1}{|T-T_\lambda|}$, le coefficient de proportionnalité étant cependant différent dans les deux cas (fig. 110). La courbe de la capacité calorifique a une forme rappelant la lettre grecque λ, et c'est ce qui détermina les appellations *point* λ et *ligne* λ. Deuxièmement, la transition He I ⇄ He II donne lieu à un saut du *coefficient de dilatation thermique*, qui est negatif pour He II.

L'hélium I se comporte comme un liquide ordinaire et possède des valeurs finies de la viscosité et de la conductibilité thermique. L'hélium II présente plusieurs propriétés spécifiques que la théorie classique ne pouvait expliquer et que la théorie quantique arriva à interpréter. C'est une autre raison de l'appeler *liquide quantique* (la première raison est sa propriété de rester liquide au zéro absolu). Ces propriétés sont la manifestation d'une seule propriété qui fut dé-

couverte en 1937 par P. L. Kapitsa (1894-1984) qui l'appela *superfluidité de l'hélium II*. Les études de l'écoulement de l'hélium II liquide à travers des conduits et des fentes étroits, effectuées antérieurement aux recherches de Kapista, conduisaient à des résultats discordants sur la viscosité de He II. Kapitsa démontra que ces désaccords étaient dus à ce que les conduits et les fentes utilisés étaient *trop larges*. Les estimations du nombre de Reynolds, fondées sur les valeurs expérimentales de la viscosité de He II, donnèrent des valeurs *très grandes*. Un nombre de Raynolds aussi grand exigerait que l'écoulement soit *turbulent*, tandis qu'on supposait qu'il était laminaire. Kapitsa arriva à un écoulement laminaire en faisant passer l'hélium II à travers des capillaires très fins (de diamètre égal et inférieur à 10^{-5} cm) et des fentes très étroites. Il constata que dans ces conditions l'hélium II liquide s'écoule à travers les capillaires *sans manifester de viscosité*. C'est ce phénomène de disparition de la viscosité qui fut appelé *superfluidité*. D'autre part, Keesom et Mac Wood ont montré en 1938 que la viscosité de He II mesurée par la méthode d'amortissement des oscillations de torsion d'un disque immergé dans He II, était finie et à proximité du point λ peu différente de celle de He I.

3. Dans ce qui suit on donne la description de plusieurs phénomènes observés dans He II, mais pour écourter l'exposé on en donnera l'interprétation fondée sur le *modèle à deux liquides* de He II, avancé en 1938 par L. Tisza (né en 1907), qui fut ensuite justifié et perfectionné par L. D. Landau en 1941. Ce modèle est fondé sur l'hypothèse que He II comporte deux constituants: un constituant normal (n) et un constituant superfluide (s), caractérisés par des densités différentes (la densité totale du liquide étant $\rho = \rho_n + \rho_s$) et des vitesses d'écoulement hydrodynamique différentes v_n et v_s. La densité totale du flux d'impulsion du liquide est $j = \rho_n v_n + \rho_s v_s$. La densité ρ_n du constituant normal tend vers zéro pour $T \to 0$ lorsque l'hélium tout entier devient superfluide; la densité ρ_s du constituant superfluide s'annule pour $T > T_\lambda$ lorsque la totalité de l'hélium se transforme en He I.

Notons que le modèle à deux liquides n'est qu'un *procédé commode de description* des phénomènes se manifestant dans He II. Comme toutes les descriptions des phénomènes quantiques en termes classiques, le modèle à deux liquides *n'est pas adéquat* si on le prend à la lettre. On ne peut dire que l'hélium II se compose d'atomes normaux et d'atomes superfluides, vu que les atomes d'hélium sont indiscernables. En permutant les positions d'un atome « normal» » et d'un atome « superfluide », à supposer que l'hélium II serait effectivement formé par ces deux espèces d'atomes, son état resterait exactement le même qu'avant permutation. Au § 61, 4 on donnera d'autres agruments prouvant qu'il ne peut être question de séparer réellement l'hélium II en des constituants normal et superfluide. D'ail-

leurs dans l'hélium II au repos on ne peut déceler aucune différence entre ces constituants. En fait il devrait être question non pas de l'existence de deux constituants mais de la *coexistence de deux mouvements, l'un normal, l'autre superfluide.* Dans ce dernier cas les densités ρ_n et ρ_s joueraient le rôle de *coefficients* indiquant la contribution de chacun de ces mouvements dans un phénomène donné. Cette remarque doit être toujours présente à l'esprit chaque fois que l'on utilisera le modèle à deux liquides.

Le mouvement normal présente toutes les caractéristiques de l'écoulement visqueux d'un liquide, tandis que le mouvement superfluide donne lieu au phénomène de superfluidité. Le modèle à deux liquides présuppose que le mouvement superfluide dans l'hélium II est non seulement dénué de viscosité mais qu'il *ne participe à aucun transport de chaleur.*

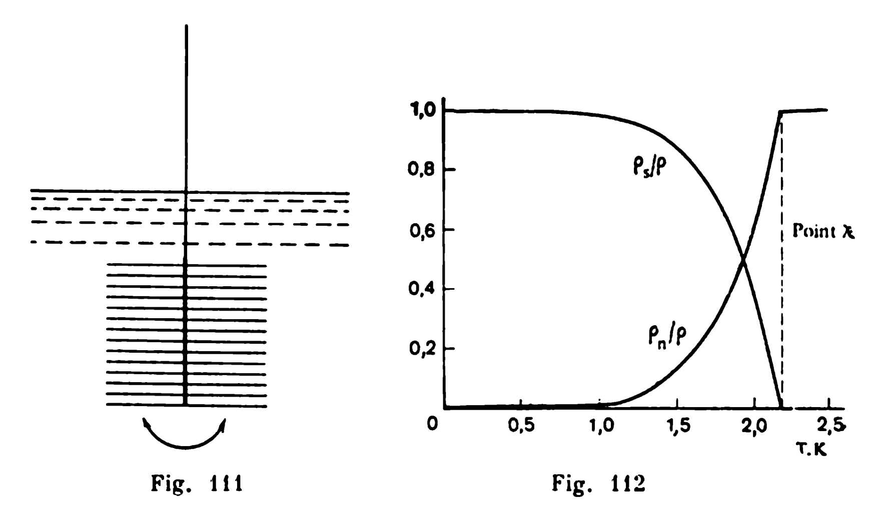


Fig. 111 Fig. 112

4. En se plaçant à ce point de vue, on comprend fort bien pourquoi la viscosité de He II disparaît lorsqu'on l'oblige à s'écouler à travers des capillaires fins. L'écoulement normal (visqueux) de He II à travers les capillaires fins ne se produit pas du tout, et seul l'écoulement superfluide, dénué de viscosité devient possible. Cet effet est d'autant plus net que le capillaire utilisé est fin, vu que dans les gros capillaires l'écoulement est turbulent.

Une expérience directe réalisée en 1946 par E. L. Andronikashvili (né en 1910) confirma le bien-fondé du modèle à deux liquides. Le schéma expérimental est illustré par la figure 111. Un empilement équidistant (0,021 cm) de disques d'aluminium (d'une épaisseur de 0,0015 cm), emmanchés sur un axe commun, était suspendu à un fil élastique et exécutait des oscillations de torsion dans un bain de He

II. Les oscillations des disques n'entraînaient que le constituant normal de He II, le consituant superfluide étant à ce point de vue absolument inactif. Lorsque $T > T_\lambda$ la totalité du liquide participait aux oscillations. Après le passage par le point λ la période des oscillations diminuait fortement, ce qui témoignait de l'apparition dans He II du constituant superfluide. A mesure que la température s'abaissait la période des oscillations continuait à diminuer. En mesurant les périodes des oscillations à différentes températures, on arrive à déterminer les teneurs relatives ρ_n/ρ et ρ_s/ρ des deux constituants d'hélium II. La figure 112, qui représente sous forme graphique les résultats de ces mesures, montre qu'au zéro absolu ne subsiste que le constituant superfluide et lorsque la température augmente sa teneur diminue. La température à laquelle le rapport ρ_s/ρ s'annule est la température de transition hélium II-hélium I. Ainsi la transition de phase dans l'hélium liquide est liée à l'apparition (ou à la disparition) du constituant superfluide. Son apparition (ou sa disparition) se produit progressivement, ce qui signifie qu'au point λ le rapport ρ_n/ρ tend vers la limite $\rho_n/\rho = 1$ *de façon continue*, sans aucun saut. C'est ce qui permet d'affirmer qu'il s'agit d'une transition de phase de second ordre.

Après ces explications on conçoit parfaitement pourquoi lors de son écoulement à travers un capillaire fin ou une fente étroite l'hélium II ne présente aucune viscosité. Cela tient à ce que seule la partie superfluide du liquide s'écoule du récipient sans aucun frottement, tandis que le constituant normal s'accumule dans le récipient, sa vitesse d'écoulement à travers le capillaire étant notablement plus petite, puisqu'elle dépend de la viscosité et du diamètre du capillaire. Le disque exécutant des oscillations de torsion dans l'hélium II est freiné par le frottement qu'il éprouve par rapport au constituant normal. C'est pour cela que l'on mesure une viscosité de valeur finie. Ainsi, dans les expériences d'écoulement à travers les capillaires on décèle le constituant superfluide de He II et dans les expériences des oscillations de torsion des disques on décèle le constituant normal.

5. Examinons maintenant quelques phénomènes qui se produisent dans l'hélium II.

Soient deux vases remplis de He II et reliés entre eux par un tube capillaire fin rempli de poudre (d'émeri, par exemple). Le constituant superfluide s'écoulera à travers les canaux sinueux de diamètres différents ($\sim$100 nm) qui se forment entre les grains de la poudre et passera d'un vase dans l'autre. Si initialement les températures de He II étaient les mêmes dans les deux vases, mais les niveaux du liquide qui y est contenu étaient différents, le constituant superfluide commence à s'écouler d'un vase dans l'autre. On a constaté que la température de He II s'élevait dans le vase d'où il s'écoule et s'abaissait dans le vase où il pénètre. C'est l'*effet* dit *mécanothermi-*

que ou *mécanocalorifique*, qui résulte de ce que le constituant superfluide de He II, qui est le seul à s'écouler d'un vase dans l'autre, ne transporte pas de chaleur. De ce fait l'énergie interne de l'hélium II contenu dans le vase d'où s'écoule le constituant superfluide reste constante, mais la quantité totale d'hélium liquide diminue. L'énergie interne revenant à l'unité de masse augmentant sans cesse, la température doit s'élever. Dans l'autre vase, où pénètrent continuellement de nouvelles portions de constituant superfluide, l'énergie interne massique de He II diminue, ce qui entraîne un abaissement de la température.

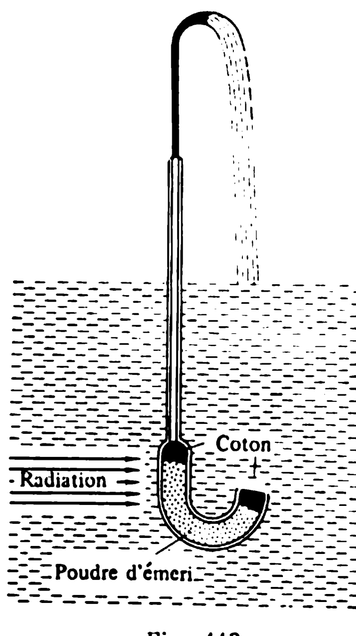


Fig. 113

6. On connaît aussi l'effet inverse appelé *effet thermomécanique*. Considérons à nouveau deux vases remplis de He II et reliés entre eux par un tube capillaire fin. A l'instant initial les températures et les niveaux de He II sont les mêmes dans les deux vases. Si nous échauffons l'un des deux vases, la teneur du constituant superfluide y diminuera et celle du constituant normal augmentera: l'équilibre thermodynamique est rompu, l'équilibre mécanique se rétablira rapidement grâce à l'écoulement du constituant superfluide du vase froid dans le vase plus chaud. Comme cela donne lieu à l'effet mécanothermique, la différence de température entre les vases s'accentuera. Comme le constituant normal ne participe pratiquement pas à l'écoulement, l'égalisation des températures s'effectue plus lentement que l'établissement de l'équilibre mécanique. Par suite tant que l'égalité des températures ne sera pas réalisée, le niveau de l'hélium dans le vase le plus froid sera plus bas que dans l'autre vase plus chaud. Ainsi, en présence d'un gradient de température dans He II, il s'établit un écoulement du constituant superfluide qui s'effectue à l'encontre de ce gradient. C'est ça l'effet thermomécanique.

Une illustration spectaculaire de l'effet thermomécanique est le jaillissement de l'hélium (*effet fontaine*) qui fut observé pour la première fois en 1938 par Allen (né en 1908) et Jones. Une des extrémités d'un tube large (fig. 113), rempli de poudre émeri tassée, était plongée dans un bain d'hélium, à son autre extrémité était fixé un tube capillaire vertical. Lorsqu'on éclairait la poudre émeri (il suffit d'utiliser la lumière d'une lampe de poche), celle-ci s'échauffait

et le constituant superfluide se précipitait du bain dans le tube si impétueusement que du tube capillaire jaillissait un jet de He II à une hauteur de 30 à 40 cm.

7. A la surface des corps immergés dans l'hélium liquide se forme une mince pellicule, ayant, à la pression de la vapeur saturante, une épaisseur de 30 nm (100 couches atomiques). Des pellicules analogues se forment aussi dans d'autres liquides (le pétrole par exemple). Or, comme He II est dénué de viscosité, son constituant superfluide s'écoule à l'intérieur de la pellicule comme il le ferait dans un

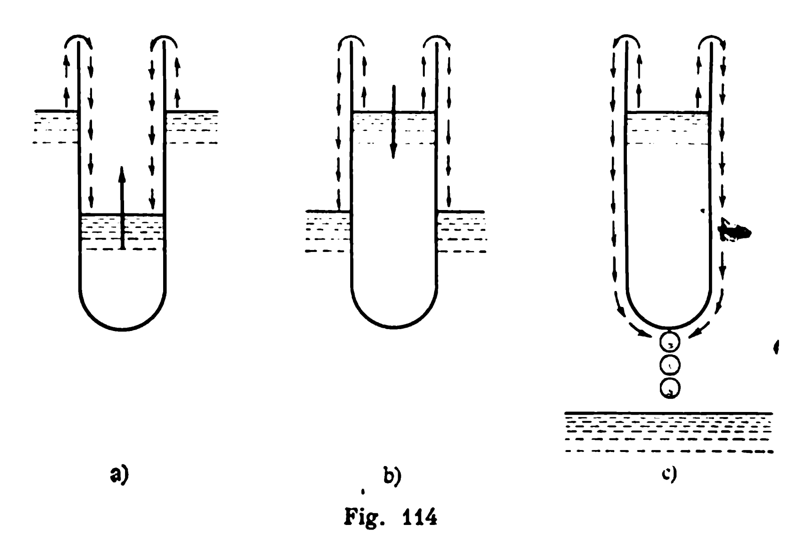


Fig. 114

canal fin; ce phénomène fut observé en 1922 par Kamerlingh Onnes. Si on enfonce une éprouvette vide dans un bain de He II, elle commence à se remplir de liquide, même si son bord supérieur se trouve au-dessus du niveau de l'hélium II (fig. 114, *a*). Le remplissage de l'éprouvette se poursuit jusqu'à ce que les niveaux du liquide dans le bain et dans l'éprouvette s'égalisent. Si après cela on relève l'éprouvette, elle se vide (fig. 114, *b*). Elle continuera à se vider, même si on la maintient toute entière au-dessus du niveau du liquide dans le bain. Dans ce dernier cas des gouttes se forment et se détachent du fond à la surface extérieure de l'éprouvette (fig. 114, *c*). Dans ces expériences la pellicule de He II se comporte comme un siphon, la force motrice mettant en mouvement le constituant superfluide étant créée, pour l'essentiel, par la différence des potentiels gravitationnels entre les deux extrémités de la pellicule. Une valeur typique de la vitesse d'écoulement du constituant superfluide de He II est de 20 cm/s environ. Le constituant normal du liquide reste pratiquement immobile dans la pellicule par suite de sa viscosité.

8. Dès le début de leurs recherches (1935) Keesom et sa fille A. Keesom constatèrent que la transformation de He I en He II s'accompagne d'un *fort accroissement* (près de 10^6 fois) *de la conductibilité thermique*. La conductibilité thermique de He II est des centaines de fois plus grande que celle des meilleurs conducteurs métalliques de la chaleur. Il s'avéra, d'autre part, que dans He II le flux de chaleur n'était pas proportionnel au gradient de température, ce qui élimine, pour He II, la notion de « coefficient de la conductibilité thermique » dans le sens usuel de ce terme. P. L. Kapitsa démontra que les anomalies du transport de la chaleur dans He II résultaient de sa superfluidité. Dans He II le transport de chaleur est assuré *par convection*. Supposons en effet que la totalité de He II se trouve au repos et qu'une différence de température s'établit entre deux points du liquide. Par suite de l'effet thermomécanique le constituant superfluide de He II se mettra alors en mouvement en allant dans le sens de la température la plus élevée. Le constituant normal de He II devra se déplacer en sens inverse afin que l'impulsion totale du liquide reste nulle. Ces deux constituants se meuvent l'un à travers l'autre sans éprouver de frottement. Comme le constituant superfluide ne transporte pas de chaleur, tout le transport de chaleur est lié au déplacement du constituant normal. C'est ce mouvement qui détermine la conductibilité thermique anomalement grande de He II. Du fait que cette dernière est anomalement grande, il est pratiquement impossible d'établir un gradient de température dans He II, et c'est ce qui explique la *brusque cessation de l'ébullition* de l'hélium liquide lorsqu'on le refroidit. Au-dessous de la température du point λ, lors du pompage, He I se comporte comme tout autre liquide et se met en ébullition. Lorsque la température s'abaisse au-dessous de celle du point λ, l'ébullition cesse brusquement. Dans un liquide ordinaire les bulles, qui déterminent la formation de la vapeur au sein du liquide, se forment lorsque la température locale dans la masse du liquide devient notablement plus grande que la température à sa surface libre. Si une telle différence de température pouvait s'établir dans He II, elle aurait disparu si rapidement que les bulles n'auraient pas eu le temps de se former. En conséquence, dans He II la vapeur ne se forme qu'à la *surface libre*, et ne pouvant se former au sein du liquide, l'ébullition de He II est impossible.

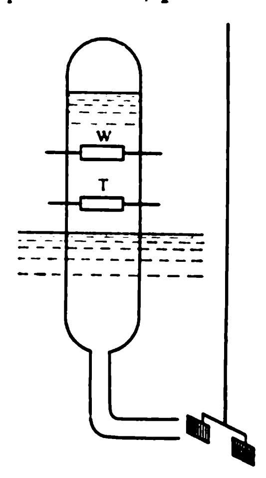


Fig. 115

9. P. L. Kapitsa observa dans des expériences l'existence dans He II des constituants superfluide et normal. Dans l'une de ses expériences (1941), dont le schéma est représenté sur la figure 115, à la

pointe inférieure d'un petit ballon de verre (ayant la forme d'un vase de Dewar en miniature) on a soudé un petit tube recourbé (longueur 6 mm, diamètre intérieur 0,5 mm). On disposait à l'intérieur du petit ballon un élément chauffant W en constantan et un thermomètre T en bronze, et on immergeait le tout dans un vase de Dewar rempli de He II. Face à l'extrémité libre du petit tube se trouvait une ailette légère attachée à un fléau suspendu à un fil de quartz, dont l'autre extrémité était équilibrée par un petit miroir. On décelait la rotation du fléau en observant le déplacement d'un rayon lumineux réfléchi par le miroir. Lorsqu'on branchait l'élément chauffant, le fléau tournait et l'ailette s'éloignait de l'extrémité du petit tube recourbé, sans que le niveau du liquide dans le ballon variât. L'explication en est la suivante. Lorsqu'on branche l'élément chauffant, le constituant superfluide de He II pénétrait dans le tube pour s'écouler dans le sens de l'élévation de la température, simultanément le constituant normal s'écoulait à travers le tube dans le sens opposé. Les deux constituants s'écoulaient en sens opposés l'un à travers l'autre et le liquide paraissait immobile dans son ensemble. Le flux normal doué de viscosité exerçait une pression sur l'ailette et la déplaçait. Le flux superfluide n'exerçait aucune pression sur l'ailette en vertu du paradoxe de D'Alembert (cf. t. I, § 100). On observait donc un phénomène curieux: un corps au repos immergé dans He II était soumis à l'action d'une force, le He II lui-même paraissant *immobile*.

10. L. Tisza a déduit du modèle à deux liquides que l'on devait observer dans He II ce que l'on appelle le *second son*. La théorie du second son a été élaborée par L. D. Landau et sa première observation expérimentale a été faite en 1946 par V. P. Pechkov (1913-1980). Dans un liquide ordinaire tout son se présente sous forme de la propagation ondulatoire des oscillations de la densité. Les ondes de ce type peuvent se propager dans He II, et on les appelle *premier son.* A $T = 0$ la vitesse du premier son dans He II est de 239 m/s environ et diminue un peu à plus haute température. Dans l'onde du premier son, à tout instant, les constituants superfluide et normal se déplacent dans le même sens, i.e. se meuvent en phase. Mais si on crée dans He II une différence de température, celle-ci est éliminée non par le mécanisme de conduction thermique usuel, mais par un mécanisme de convection rapide lorsque les deux constituants se déplacent en sens opposés. Si on excite dans He II des oscillations de température, les déplacements en sens opposés des constituants superfluide et normal font que la densité totale de He II reste constante et que le flux d'impulsion total $j = \rho_s v_s + \rho_n v_n$ est égal à zéro. Les deux constituants oscillent alors *en contrephase*, mais le transport de chaleur est assuré par le seul constituant normal. Ainsi, dans le cas considéré apparaît une *onde de température* dans laquelle varie la température et non pas la densité. C'est ce que l'on appelle le second son.

La vitesse du second son dans He II dépend fortement de la température. Pour $T = 0$ sa vitesse est proche de 139 m/s et s'annule pour $T = T_\lambda$. Selon la théorie de Landau, à $T = 0$, la vitesse du second son est, en accord avec l'expérience, $\sqrt{3}$ fois plus petite que la vitesse du premier son.

PROBLÈMES

Démontrer que la quantité de chaleur qu'il faut fournir à l'hélium pour porter sa température de t à $t + dt$ tend vers zéro lorsque $dt \to 0$, même si l'intervalle considéré inclut le point λ, où la capacité calorifique de l'hélium devient infinie.

§ 61. Notion de la théorie de superfluidité

1. Le modèle à deux liquides ne convient évidemment qu'à la *description* des phénomènes, sans pouvoir les interpréter. La nature du phénomène fut expliquée en 1941 par L. D. Landau qui justifia simultanément le modèle à deux liquides. Dans ce qui suit on présente quelques notions qualitatives de la théorie de Landau. Cette théorie postule tout d'abord qu'au zéro absolu l'hélium II se trouve dans son état quantique *fondamental*, donc l'état le plus bas (non excité). Si le He II se trouve dans cet état, le déplacement d'un corps dans ce liquide ne s'accompagne d'aucun frottement. Pour faire apparaître un frottement il faut que le déplacement du corps dans He II le fasse passer à l'état *excité*. Comme l'interaction mutuelle des atomes d'hélium liquide est forte, les états excités doivent être *collectifs* (concernant la totalité de l'hélium) au lieu de correspondre à une excitation d'atomes individuels. Les excitations de ce type sont de nature quantique. Les quanta d'excitation de He II sont appelés *excitations élémentaires* et sont caractérisés par des valeurs déterminées de l'impulsion et de l'énergie, i.e. se comportent comme des quasi-particules. Le caractère collectif des perturbations élémentaires est un trait essentiel de la théorie de Landau.

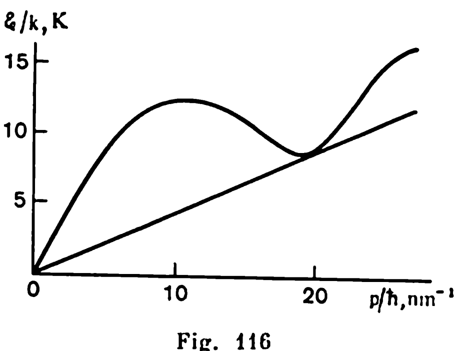


Fig. 116

La théorie pour être valable doit préciser en premier lieu la *loi de dispersion* ou le spectre énergétique des excitations élémentaires, i.e. la dépendance de l'énergie $\mathscr{E}$ avec l'impulsion p ou plus exactement avec le module p de l'impulsion, puisque l'hélium liquide est un milieu isotrope. Ce premier problème théorique a été traité par Landau sous forme de *postulats* en se fondant sur des données empiri-

ques. La courbe $\mathscr{E} = \mathscr{E}(p)$ qui fut finalement établie par Landau en 1947 est représentée sur la figure 116. Sur la partie initiale de la courbe l'énergie $\mathscr{E}$ croît linéairement avec l'impulsion: $\mathscr{E} = c_{son}p$, donc exactement comme dans le cas de l'excitation d'ondes acoustiques dans un réseau cristallin. Sur cette partie de la courbe, il se produit dans He II une excitation de *phonons*, la constante c_{son} étant la vitesse de phase du son dans He II. Cette forme du spectre phononique des excitations élémentaires prédétermine l'allure de la variation de la capacité calorifique de He II en fonction de la température à proximité du zéro absolu : comme pour les corps solides, la capacité calorifique est proportionnelle au cube de la température T.

Lorsque p augmente la courbe $\mathscr{E} = \mathscr{E}(p)$ s'incurve vers le bas et passe par un minimum au point p_0. A proximité du minimum l'énergie peut être approchée par la formule

$$\mathscr{E} = \Delta + \frac{(p-p_0)^2}{2m^*}, \quad (61.1)$$

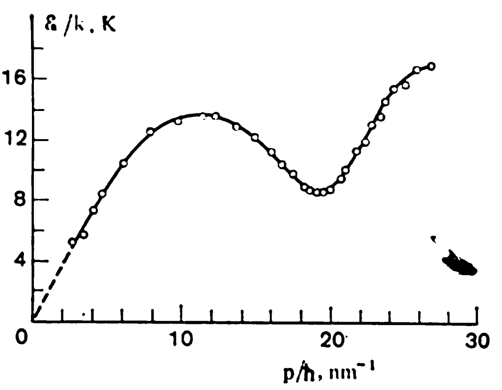


Fig. 117

la constante m^* jouant le rôle de *masse efficace*. Les excitations élémentaires correspondant à cette partie de la courbe sont appelées *rotons*.

Il importe de noter que la loi de dispersion $\mathscr{E} = \mathscr{E}(p)$ fait toujours intervenir l'*impulsion vraie* et non pas la quasi-impulsion, comme c'est le cas pour le réseau cristallin. La raison en est que l'hélium liquide est un *milieu isotrope* denué de structure cristalline.

Ultérieurement la courbe $\mathscr{E} = \mathscr{E}(p)$ postulée par Landau fut confirmée par l'expérience. On étudiait la diffusion d'un faisceau de neutrons monochromatiques dans He II. On peut considérer que ce processus consiste en des chocs entre les neutrons et les quasi-particules existant dans He II, lors desquels les lois de conservation de l'énergie et de l'impulsion sont respectées. Connaissant l'énergie initiale et la direction des neutrons incidents et mesurant l'énergie des neutrons diffusés sous différents angles de diffusion (par rapport à la direction initiale des neutrons), on peut calculer l'énergie et l'impulsion des quasi-particules, donc déterminer tout le spectre énergétique des excitations élémentaires. La courbe expérimentale $\mathscr{E} = \mathscr{E}(p)$ ainsi obtenue est représentée sur la figure 117. Les valeurs numériques des constantes (extrapolées à la pression nulle pour une densité $\rho = 0{,}145$ g/cm³) déduites de ces expériences ainsi que de calculs

thermodynamiques sont les suivantes:

$$c_{son} = 2,4\cdot 10^4 \text{ cm/s}, \quad \Delta/k = 8,7 \text{ K}, \\ p_0/\hbar = 1,9\cdot 10^8 \text{ cm}^{-1}, \quad m^* = 0,16 m_{He}. \tag{61.2}$$

2. Quoique corroboré par les résultats expérimentaux, le fait que le spectre énergétique ait été postulé dans la théorie de Landau, déçoit. Il est cependent impossible de déterminer ce spectre par des calculs pour un liquide tel He II dont les particules interagissent fortement. Plus tard, en 1946, N. N. Bogolioubov (né en 1908) a réussi à traiter ce problème pour un gaz de bosons *à interactions mutuelles faibles* (il a été d'ailleurs amené à supposer que les forces de répulsion s'exerçant entre ces particules étaient supérieures aux forces d'attraction). Bogolioubov trouva que pour les faibles valeurs de p, le spectre des excitations élémentaires est de type phononique, comme le postule la théorie de Landau.

On ne peut exposer ici la théorie de Bogolioubov, mais on notera qu'il a dû admettre que le gaz se compose de *bosons* et *non pas de fermions* (comme dans toutes les théories de la superfluidité). Dans le cas de bosons il se produit une condensation de Bose-Einstein (chose impossible pour les fermions en vertu du principe de Pauli). Cette condensation (voir t. II, § 82) résulte de ce qu'à partir d'une certaine température T_0, dite *température de dégénérescence*, les particules du gaz de Bose commencent à s'accumuler sur le niveau énergétique le plus bas; au zéro absolu toutes les particules occupent ce niveau. Les particules de Bose qui se sont accumulées sur le niveau énergétique le plus bas forment un *condensat*. Bogolioubov arriva à démontrer que, moyennant certaines hypothèses sur la nature des forces d'interaction entre les particules, la formation d'un condensat, dans un gaz de bosons peu différent d'un gaz parfait, pouvait être assimilée à une transition de phase de second ordre et que le mouvement de ce condensat jouissait de propriétés superfluides. L'analogie entre ce phénomène et la superfluidité observée de He II liquide se manifeste en ce que le calcul de la température de dégénérescence du gaz de bosons fournit la valeur de $T_0 = 3,14$ K peu différente de la température de la transition de phase de He I en He II ($T_\lambda = 2,17$ K). Néanmoins les résultats de Bogolioubov concernent un *gaz* de particules à interactions mutuelles faibles et non pas l'hélium II *liquide* où les interactions sont fortes. Il serait souhaitable d'élaborer sur les mêmes bases une théorie de la superfluidité pour les liquides, mais jusqu'à présent on n'a pas réussi à le faire.

3. Démontrons maintenant que le spectre énergétique des excitations élémentaires postulé par Landau implique la superfluidité du liquide. Considérons d'abord He II au zéro absolu, lorsqu'il ne comporte pas d'excitations élémentaires. Supposons que ce liquide s'écoule à travers un tube capillaire avec une vitesse constante v, et cher-

chons à déterminer à quelles vitesses v les excitations élémentaires n'apparaissent pas encore et l'hélium n'éprouve aucun frottement et à quelles autres vitesses les excitations apparaissent et déterminent le frottement du liquide contre les parois du tube. Pour faciliter la discussion il est commode d'utiliser un référentiel dans lequel le liquide est au repos et le tube se déplace avec la vitesse $-v$. La raison du choix de ce référentiel tient à ce que la loi de dispersion concerne justement le référentiel où *le liquide est au repos* (fig. 117).

Supposons maintenant qu'une excitation élémentaire, i.e. une quasi-particule d'énergie $\mathcal{E}$ et d'impulsion p, est apparue dans le liquide. Par effet de recul, la vitesse du tube varie et devient égale à $\mathcal{E} - v_1$. En appliquant la loi de la conservation de l'énergie, on peut écrire

$$\mathcal{E}+\frac{M}{2}v_1^2-\frac{M}{2}v^2=0,$$

où M est la masse du tube capillaire. Comme elle est grande, la variation de la vitesse du tube est très petite, ce qui permet de remplacer la somme $v + v_1$ par $2v$ et d'obtenir

$$\mathcal{E}=\frac{M}{2}(v-v_1)(v+v_1)=Mv(v-v_1).$$

En vertu de la loi de la conservation de l'impulsion

$$p+Mv_1=Mv,$$

d'où

$$\mathcal{E}=pv. \tag{61.3}$$

On peut en conclure que $\mathcal{E} \leqslant pv$, soit $v \geqslant \mathcal{E}/p$. Ainsi, pour qu'apparaisse une excitation élémentaire il faut que la vitesse v du liquide soit supérieure à la valeur minimale de $\mathcal{E}/p$

$$v \geqslant \min(\mathcal{E}/p). \tag{61.4}$$

Sur la courbe de dispersion le minimum de $\mathcal{E}/p$ correspond au point où la dérivée $\frac{d}{dp}(\mathcal{E}/p)$ s'annule, i.e. au point

$$\frac{d\mathcal{E}}{dp}=\frac{\mathcal{E}}{p}, \tag{61.5}$$

où la droite $\mathcal{E}/p = \text{const}$ passant par l'origine des coordonnées est tangente à la courbe de dispersion $\mathcal{E} = \mathcal{E}(p)$. Sur la branche phononique $\mathcal{E} = c_{\text{son}}\, p$, ce qui implique que les phonons ne seront excités que si la vitesse $v \geqslant c_{\text{son}}$.

Ce résultat n'est pas inattendu. On sait en effet que lorsqu'un corps est animé d'un mouvement uniforme au sein d'un liquide il commence à générer des *ondes acoustiques de Mach* (1838-1916) dès

que la vitesse du corps devient supérieure à la vitesse de phase du son dans le fluide. Un effet analogue apparaît lorsqu'une charge se déplace dans un milieu avec une vitesse supérieure à la vitesse de phase de la lumière dans la phase considérée (*effet Vavilov-Cerenkov*). Par conséquent, la vitesse du son joue le rôle de *vitesse critique* v_{cr}, en ce sens qu'un corps mobile ne produit un son que lorsque sa vitesse devient supérieure à la vitesse critique.

Pour déterminer la vitesse critique qui donnera naissance aux rotons de la théorie de Landau, il suffit de faire passer par l'origine des coordonnées (fig. 117) une tangente à la branche rotonique de la courbe $\mathcal{E} = \mathcal{E}(p)$. Cette vitesse critique est évidemment égale à $v_{cr} = \mathcal{E}/p$, où $\mathcal{E}$ et p sont l'énergie et l'impulsion du roton au point de contact. On trouve pour les rotons $v_{cr} = 60$ m/s.

4. Considérons maintenant le cas où la température du liquide s'écoulant à travers le tube est différente, quoique proche du zéro absolu. Dans ces conditions, même dans un liquide au repos peuvent apparaître de faibles excitations qui se comportent comme un *gaz parfait de quasi-particules indépendantes*. La formule (61.4), établie plus haut, reste valable puisqu'elle exprime la condition de formation des excitations élémentaires par suite du mouvement du liquide par rapport aux parois du tube capillaire, et que pour l'établir on n'avait pas imposé que la température du liquide soit égale au zéro absolu.

Dans un liquide au repos, se trouvant à une température $T \neq 0$, les quasi-particules qui s'y trouvent sont animées de mouvements désordonnés, de sorte qu'en moyenne le gaz de quasi-particules ne transporte aucune impulsion. Mais si le liquide se déplace, il apparaît une direction privilégiée de déplacement des quasi-particules (dans le sens d'écoulement du liquide) et l'impulsion moyenne du gaz de quasi-particules devient différente de zéro et se trouve transportée dans le même sens. Lors des chocs des quasi-particules avec les parois du tube, une partie au moins de cette impulsion est transmise aux parois, ce qui fait apparaître un frottement. En conséquence, la partie du liquide transportant le gaz d'excitations élémentaires s'écoule à travers le tube capillaire comme un liquide ordinaire doué de viscosité. La partie restante du liquide est superfluide, puisque pour une vitesse d'écoulement inférieure à la vitesse critique, son mouvement par rapport aux parois du tube capillaire ne produit pas de nouvelles quasi-particules. Cette partie du liquide n'échange ni énergie ni impulsion avec les parois du tube capillaire ou avec le gaz de quasi-particules. C'est par ces considérations que la théorie de Landau justifie les conceptions introduites dans le modèle à deux liquides. Mais la théorie de Landau démontre clairement qu'il est impossible de séparer réellement les constituants superfluide et normal, puisque les excitations élémentaires sont *collectives* et non pas des excitations de particules individuelles.

5. La valeur de la vitesse critique d'écoulement ($v_{cr} \approx 60$ m/s) calculée par la théorie de Landau à partir de la courbe de dispersion ne coïncide pas avec les résultats des mesures. L'expérience montre que la vitesse critique dépend fortement de la température et du rayon du tube capillaire ou de la largeur de la fente à travers lesquels s'écoule le He II. La valeur expérimentale de la vitesse critique est de deux à trois ordres plus petite que la valeur théorique. On doit en inférer que la théorie de Landau est insuffisante, car elle n'a pas tenu compte de tous les facteurs. En réalité, la théorie de Landau n'indique que la *limite supérieure* de la vitesse d'écoulement, au-delà de laquelle le phénomène de superfluidité doit sûrement disparaître.

On ne commence à comprendre les causes de ce désaccord entre théorie et expérience qu'à la suite de l'étude des propriétés de He II animé d'un mouvement de rotation. Selon le modèle à deux liquides, lors de la rotation d'un vase cylindrique rempli de He II, la partie superfluide du liquide ne devrait pas être entraînée dans la rotation. Or les expériences ont montré que si la vitesse angulaire de rotation du cylindre était suffisamment grande, la *partie superfluide* est entraînée au même titre que la partie normale du liquide. Néanmoins dans He II mis en rotation dans son ensemble on observait des phénomènes de superfluidité aussi caractéristiques que l'effet thermomécanique et la propagation du second son.

Pour expliquer ces résultats, Onsager (né en 1903) avança en 1949 l'hypothèse, qui fut confirmée plus tard par l'expérience, que du fait de la *nature quantique* de l'état superfluide de He II, des rotations peuvent y être excitées, mais elles doivent être quantifiées. Plus tard (en 1955) cette hypothèse fut reprise par Feynman (né en 1918). Si un atome d'hélium tourne autour de l'axe du cylindre sur une trajectoire circulaire de rayon r, son moment cinétique doit être égal à $n\hbar$, i.e.

$$m_{\text{He}} r v_s = n\hbar,$$

où v_s est la vitesse de l'écoulement superfluide et n un nombre entier. Cette formule n'est valable que pour des distances r notablement plus grandes que la distance interatomique moyenne dans ^{4}He, i.e. $4 \cdot 10^{-8}$ cm environ, puisque ce n'est qu'à ces distances que l'hélium peut être assimilé à un liquide continu. Le rayon r du cercle peut prendre de très grandes valeurs jusqu'au rayon du cylindre, de sorte qu'il s'agit de la *quantification de rotations macroscopiques*.

La ligne autour de laquelle se produit la rotation de He II est appelée *ligne tourbillonnaire*. Dans le cas d'une rotation autour de l'axe du cylindre cette ligne coïncide évidemment avec cet axe et s'étend du fond du cylindre à la surface du liquide. Le fl ide qui tourne autour d'une ligne tourbillonnaire est appelé *tourbillon quantique*; celui-ci ne peut apparaître que pour une certaine vitesse minima-

le de rotation du cylindre Ω_{cr}, car sa formation implique une dépense d'énergie. En fait n'apparaissent que des tourbillons avec $n = 1$, car l'énergie cinétique étant proportionnelle à $v_s^2 \sim n^2$, il est énergétiquement plus avantageux de former deux tourbillons quantiques avec $n = 1$ qu'un seul avec $n = 2$. Par conséquent, la vitesse de déplacement v_s du liquide superfluide lors de sa rotation autour d'une ligne tourbillonnaire est donnée par

$$v_s = \frac{\hbar}{m_{He} r}. \tag{61.6}$$

Lorsque la vitesse de rotation du cylindre atteint sa valeur critique Ω_{cr}, il ne se forme qu'un seul tourbillon quantique dont la ligne coïncide avec l'axe du cylindre. Lorsque Ω croît le nombre de tourbillons croît, et leurs lignes sont parallèles à l'axe du cylindre. Lorsque Ω est grand les lignes tourbillonnaires transpercent presque uniformément tout le volume de l'hélium liquide.

Les excitations élémentaires correspondant à la partie normale de He II peuvent entrer en collision avec les tourbillons quantiques de la partie superfluide et échanger de l'énergie et de l'impulsion avec ces derniers. C'est ce qui détermine la mise en rotation de la partie superfluide lors de la rotation du cylindre.

6. Lorsque la vitesse d'écoulement de He II à travers un tube capillaire devient grande les lignes tourbillonnaires peuvent s'incurver et même se fermer sur elles-mêmes en formant les *anneaux tourbillonnaires*. Ces derniers possèdent une énergie et une impulsion déterminées et peuvent se déplacer dans tout le volume de l'hélium liquide. On pourrait dire que les anneaux tourbillonnaires sont des *excitations élémentaires* particulières qui apparaissent dès que la vitesse d'écoulement de l'hélium superfluide devient supérieure à une certaine vitesse critique v_{cr}. La vitesse critique v_{cr} dépend, d'une part, de l'énergie qui doit être mise en jeu pour former un anneau quantique tourbillonnaire et, d'autre part, des conditions d'écoulement de He II, notamment du rayon de tube capillaire; lorsque ce rayon augmente la vitesse critique diminue rapidement. Pour des rayons du capillaire de l'ordre de 10^{-5} cm, rayons pour lesquels on observe le phénomène de superfluidité, la vitesse critique est notablement plus petite que la limite théorique supérieure prévue par la formule de Landau (61.4). Lorsque la vitesse d'écoulement v est plus petite que la vitesse critique les anneaux tourbillonnaires quantiques ne se forment pas et on observe la superfluidité. Losrque la vitesse d'écoulement est supérieure à v_{cr} (quoique reste inférieure à la limite supérieure prévue par la théorie de Landau) l'écoulement superfluide de He II devient instable vis-à-vis de la formation des anneaux tourbillonnaires quantiques. Les anneaux tourbillonnaires quantiques qui apparaissent dans le constituant superfluide de He II peuvent échanger de l'énergie et de l'impulsion avec les quasi-particules du

constituant normal, ce qui fait apparaître un frottement entre ces constituants et conduit à la disparition de la superfluidité.

7. Examinons brièvement la question de la superfluidité de $^3_2$He. L'atome de $^3_2$He contient un nombre impair de nucléons (deux protons et un neutron) et un nombre pair (2) d'électrons. Comme chacune de ces particules possède un spin 1/2, le spin de l'atome $^3_2$He est nécessairement demi-entier, de sorte que les atomes $^3_2$He obéissent à la statistique de Fermi-Dirac. A ce point de vue ils se distinguent des atomes $^4_2$He qui ont un spin entier et de ce fait se conforment à la statistique de Bose-Einstein. Seuls les systèmes dont les particules se conforment à la statistique de Bose-Einstein peuvent former un condensat et manifester la superfluidité. Aussi, pensa-t-on tout d'abord que le $^3_2$He ne pouvait devenir superfluide mais on changea d'avis en 1974 lorsque fut élaborée la théorie de la supraconductibilité (cf. § 62). En effet, puisque des forces d'attraction de Van der Waals s'exercent entre les atomes de $^3_2$He et que, à grande distance, ces forces décroissent en raison inverse de la distance élevée à la puissance sept, il devient énergétiquement avantageux que les atomes $^3_2$He s'accouplent et c'est ce qui se produit aux très basses températures. Le spin de ces paires d'atomes est toujours entier, ce qui fait que les paires sont régies par la statistique de Bose-Einstein et qu'on peut donc s'attendre à l'apparition de la superfluidité chez $^3_2$He. C'est ce qui fut confirmé en 1974. La superfluidité de $^3_2$He fut décelée au-dessous de 0,00265 K (sous une pression de 34 atm environ). Si on diminue la pression extérieure jusqu'à 21 atm, la température de transition de $^3_2$He liquide à l'état superfluide s'abaisse jusqu'à 0,0024 K.

8. Pour conclure ce paragraphe, examinons la possibilité d'apparition dans l'hydrogène de phénomènes quantiques macroscopiques tels que la superfluidité. Le comportement quantique de l'hélium, notamment son aptitude à rester liquide à des températures proches du zéro absolu, est essentiellement déterminé par les *vibrations de zéro*. Les masses des atomes et des molécules d'hydrogène étant plus petites que celle de l'atome d'hélium, leurs vibrations de zéro sont donc plus intenses. D'autre part, l'atome d'hydrogène se compose de deux particules (proton et électron) ayant chacune un spin 1/2 et se comporte donc comme un *boson*. La molécule d'hydrogène est aussi un *boson*. On pourrait donc s'attendre à ce que, à proximité du zéro absolu, l'hydrogène devienne superfluide. Or, le gaz d'hydrogène biatomique, quoique manifestant certaines propriétés quantiques, au refroidissement se liquéfie d'abord, puis se solidifie à proximité du zéro absolu.

La raison en est que les particules d'hydrogène *interagissent entre elles*. Les forces chimiques homéopolaires s'exerçant entre les atomes déterminent la formation d'un état lié — les molécules H_2, entre lesquelles s'exercent des forces notablement plus grandes que celles

qui s'exercent entre les atomes symétriques et inactifs d'hélium.

On doit cependant tenir compte de ce que les forces d'interaction s'exerçant entre les atomes d'hydrogène neutres se trouvant à des distances supérieures à une certaine valeur, ne peuvent être des forces d'attraction mutuelle que si les spins électroniques des atomes sont *antiparallèles* (état singulet). Si les spins des électrons des atomes d'hydrogène sont *parallèles* (état triplet), les forces qui s'exercent entre les atomes sont toujours des forces de répulsion, et la formation de molécules devient impossible (cf. § 50).

On peut cependant obtenir de l'hydrogène atomique dont tous les atomes ont des spins parallèles, ce qui exclut la formation de molécules. Pour cela on soumet l'hydrogène dissocié à l'action d'un champ magnétique fort ($\sim 10^5$ Gs) très inhomogène. Dans ces conditions les atomes à spins antiparallèles seront spatialement séparés, comme cela se produit dans les expériences de Stern et Gerlach. On doit tenir compte de ce que le moment magnétique du noyau (du proton) est près de mille fois plus petit que celui de l'électron, ce qui permet de négliger la force qu'exerce le champ magnétique sur le noyau devant la force qu'exerce le même champ sur l'électron.

Cette méthode d'obtention d'hydrogène atomique à spins électroniques parallèles à été utilisée en 1980 par les physiciens hollandais Silvera et Walraven. Pour mener à bien leur entreprise ils durent surmonter de nombreuses difficultés expérimentales qu'on omettra de signaler ici. En théorie on peut s'attendre à ce que l'hydrogène atomique ainsi préparé reste gazeux à proximité du zéro absolu, sans même se liquéfier. Ces physiciens réussirent à maintenir le gaz à l'état atomique pendant plusieurs heures à très basse température en présence d'un champ magnétique fort. L'hydrogène atomique restait gazeux jusqu'à 0,08 K, la plus basse température qui put être atteinte dans ces expériences.

§ 62. Eléments de la théorie de la supraconductibilité

1. Avant de s'attaquer à ce paragraphe le lecteur est invité à relire le § 80 du t. III, où sont exposées les principales données expérimentales concernant la supraconductibilité ainsi que l'histoire de la découverte de ce phénomène. Dans ce paragraphe on présente quelques notions de la théorie de la supraconductibilité, dans la mesure compatible avec un cours de physique générale.

Après que fut élaborée la théorie de la superfluidité on en vint tout naturellement à concevoir la supraconductibilité comme la superfluidité du « fluide électronique » constitué par les électrons de valence des atomes du métal. Cette idée juste se heurte cependant à de grosses difficultés. Comme les électrons ont un spin égal à 1/2, ils doivent satisfaire au principe de Pauli et à la statistique de

Fermi-Dirac, ce qui implique que les électrons sont des fermions. Par contre, le principe de Pauli ne s'applique pas aux atomes $^4_2$He, puisque ces atomes possèdent un spin entier. A la différence des électrons, les atomes $^4_2$He sont régis par la statistique de Bose-Einstein et sont donc des bosons. C'est ce qui détermine une différence essentielle entre les états fondamentaux des atomes de $^4_2$He et des électrons. En particulier, les atomes de $^4_2$He sont susceptibles de subir une condensation de Bose-Einstein, ce qui, à première vue, semble exclu dans le cas des électrons.

2. La découverte de l'*effet isotopique*, faite en 1950 au cours d'expériences avec les isotopes de mercure et d'étain, contribua largement à surmonter ces difficultés. L'effet isotopique s'observe chez les supraconducteurs purs, possédant plusieurs isotopes. Il s'avéra que les réseaux cristallins construits avec différents isotopes d'un seul et même élément chimique vérifient la relation $T_{cr}\sqrt{M} = C =$ const. Ici T_{cr} est la température critique de transition à l'état supraconducteur et M le nombre massique des atomes formant le réseau cristallin. La constante C a la même valeur pour tous les isotopes d'un élément donné, mais des valeurs différentes pour différents éléments. Or, comme la fréquence des vibrations du réseau dépend de M $\omega \sim \sqrt{1/M}$, l'existence de l'effet isotopique induit à penser que la supraconductibilité serait déterminée par les interactions des électrons avec les vibrations du réseau cristallin.

Cette conclusion est encore renforcée par le fait que les métaux purs de grande conductibilité (notamment les éléments du premier groupe du système périodique: Li, Na, K, Cu, Ag, Au) ne manifestent pas de propriétés supraconductrices. La raison en est que les métaux bons conducteurs possèdent une grande longueur de libre parcours des électrons, et comme celle-ci dépend de leurs interactions avec le réseau cristallin, on en conclut que dans ces métaux ces interactions sont faibles. Par contre chez les métaux mauvais conducteurs, la longueur de libre parcours des électrons est petite, ce qui signifie que leurs interactions avec les vibrations du réseau sont fortes; or c'est ça qu'il faudrait pour que le métal puisse passer à l'état supraconducteur.

La découverte de l'effet isotopique indiqua dans quelle voie devait-on chercher à établir une théorie microscopique de la supraconduction. Un premier pas fut franchi en 1950 par Fröhlich (né en 1905) qui avait prévu, en théorie, l'effet isotopique, pratiquement en même temps que cet effet était découvert expérimentalement. A la suite d'une étude fondamentale de Cooper (né en 1930), une théorie microscopique conséquente de la supraconduction fut élaborée en 1957 par Bardeen (né en 1908), Cooper et Schrieffer (né en 1931). D'après les initiales de ses inventeurs cette théorie s'appelle brièvement théorie de B.C.S. Une méthode plus élaborée de la théorie de la supraconduction a été conçue par N. N. Bogolioubov à peu près

à la même époque. Les prévisions de Cooper relatives à l'accouplement des électrons (voir sous-paragraphe 4 ci-après) ont permis à Bogolioubov de développer la méthode qu'il élabora auparavant en théorie de la superfluidité. Une autre méthode encore dans la théorie de la supraconduction fut élaborée par L. P. Gorkov (né en 1929) et Nambu (né en 1921).

3. Il est bien connu que les vibrations du réseau cristallin sont quantifiées et peuvent être assimilées à de quasi-particules appelées phonons. Dans un métal, les électrons de conduction peuvent échanger des phonons à tout instant; un électron émet un phonon et un autre le capte; ces phonons sont dits *virtuels*, leur durée de vie étant très courte, leur énergie n'est pas fixée mais satisfait au principe d'indétermination temps-énergie. Cela permet de ne pas tenir compte de la loi de la conservation de l'énergie dans les processus d'interaction d'un phonon virtuel avec un électron. L'électron ayant émis un phonon virtuel subit un recul, autrement dit son impulsion change. L'impulsion de l'électron qui a capté le phonon change également. Tout se passe comme s'il existait entre les électrons, outre la répulsion coulombienne, une interaction supplémentaire. Cette dernière peut donner lieu à une attraction ou à une répulsion mutuelle des électrons.

Cet effet peut être illustré par une analogie classique. Considérons deux patineurs qui se renvoient tout le temps une balle. Par suite du recul qu'ils subissent à chaque coup donné sur la balle il s'établit entre eux une répulsion. Si on remplaçait la balle par un boomerang, on aurait, au lieu d'une répulsion, une attraction. Pour cela les deux patineurs doivent se tourner le dos et lancer le boomerang dans une direction opposée à celle de son partenaire. Bien entendu, le comportement du boomerang est déterminé par le milieu atmosphérique dans lequel il évolue. Dans le cas des phonons le rôle du milieu atmosphérique est assumé par le réseau cristallin.

4. Dans la théorie des métaux développée par Bloch (cf. § 59) l'état des électrons est caractérisé par des fonctions d'onde *monoélectroniques*, ce qui implique que les électrons sont supposés indépendants. Le phénomène de la supraconduction ne se laisse pas interpréter dans le cadre de ce schéma théorique. Par suite des interactions entre les électrons et les phonons, il existe entre les électrons des corrélations dont on doit tenir compte. Dans la théorie de B.C.S. on ne tient compte que des *corrélations paires*, ce qui s'avéra suffisant pour donner une explication, ne serait-ce que qualitative, des principales propriétés des supraconducteurs. Les forces d'attraction phononiques s'exerçant entre une paire d'électrons correlés doivent être supérieures à la force de répulsion coulombienne s'exerçant entre ces mêmes électrons, afin qu'une attraction s'exerce entre ces derniers. On dit couramment que ces électrons couplés forment une *paire de Cooper* (en l'honneur de Cooper qui a prévu ce couplage en 1956). Les paires

de Cooper étant des bosons, l'interdiction de Pauli ne les concerne pas et elles peuvent donc exécuter des mouvements superfluides.

Pour mettre en évidence les propriétés des paires de Cooper, supposons d'abord que le métal se trouve à la température du zéro absolu, et qu'il ne soit pas parcouru par un courant électrique. Comme les électrons obéissent au principe de Pauli, chacun des états quantiques ne peut être occupé que par un seul électron. De ce fait tous les états quantiques situés au-dessous d'une certaine énergie $\mathcal{E}_F$, appelée *énergie de Fermi*, seront occupés et tous les états correspondant aux énergies supérieures à $\mathcal{E}_F$ seront vacants. L'énergie $\mathcal{E}$ est fonction de l'impulsion: $\mathcal{E} = \mathcal{E}(p)$. En particulier, dans l'espace des impulsions l'équation $\mathcal{E}_F = \mathcal{E}(p)$ définit une certaine surface dite *surface de Fermi*. Pour simplifier, posons que le métal est isotrope; dans ce cas la surface de Fermi dégénère en une sphère $\mathcal{E}_F = p^2/2m$ de rayon p_F. Si la température du métal augmente tout en restant plus petite que T_{cr}, les électrons peuvent s'échapper de la sphère de Fermi et venir occuper des états quantiques de plus grande énergie. De ce fait de part et d'autre de la sphère de Fermi il se forme une couche mince délimitée par des sphères de rayons $p_F - \Delta p/2$ et $p_F + \Delta p/2$, qui n'est que partiellement occupée par les électrons. A proximité de la température critique T_{cr}, les énergies des électrons se trouvant sur les bords extrêmes de cette couche se distinguent d'une quantité de l'ordre de kT_{cr}. Un calcul simple montre que cette quantité est 10^4 à 10^5 fois plus petite que l'énergie cinétique moyenne revenant dans un métal à un électron de valence satisfaisant à la distribution de Fermi (cf. t. III, § 99).

Supposons maintenant que les deux électrons d'une paire de Cooper, d'impulsions $\boldsymbol{p}_1$ et $\boldsymbol{p}_2$ entrent en collision et que leurs nouveaux états soient caractérisés par les impulsions $\boldsymbol{p}_1'$ et $\boldsymbol{p}_2'$. De toute évidence, avant la collision les états d'impulsions $\boldsymbol{p}_1'$ et $\boldsymbol{p}_2'$ doivent être vacants, i.e. doivent se situer dans la mince couche sphérique invoquée plus haut. Il en va de même pour les vecteurs initiaux $\boldsymbol{p}_1$ et $\boldsymbol{p}_2$, car autrement après le choc les états d'impulsions $\boldsymbol{p}_1$ et $\boldsymbol{p}_2$ resteraient vacants, ce qui serait à exclure si les états se trouvaient dans la partie de l'espace des impulsions délimitée par la sphère interne de la couche. En vertu de la loi de la conservation de l'impulsion on doit avoir $\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_1' + \boldsymbol{p}_2' = \vec{\mathcal{P}} = \text{const}$, où $\vec{\mathcal{P}}$ est l'impulsion résultante des électrons qui reste constante dans les chocs. Le processus de collision est illustré par la figure 118. L'origine du vecteur $\boldsymbol{p}_1$ et l'extrémité du vecteur $\boldsymbol{p}_2$ doivent se trouver dans le rectangle hachuré du haut, qui est formé par l'intersection de deux couches sphériques de rayons moyens p_1 et p_1'. De même, le rectangle hachuré du bas doit contenir l'origine du vecteur $\boldsymbol{p}_1'$ et l'extrémité du vecteur $\boldsymbol{p}_2'$. Il tombe sous le sens que les chocs seront d'autant plus fréquents que le vecteur $\vec{\mathcal{P}}$ sera court. Pratiquement, les chocs ne pourront se produire que si les

deux couches sphériques coïncident, i.e. lorsque $\vec{\mathcal{P}} = 0$. On doit donc avoir $p_1 = -p_2$, $p'_1 = -p'_2$. Ainsi, en l'absence de courant les impulsions des électrons appartenant à une paire de Cooper doivent être de même module et de sens opposé.

On a déjà indiqué qu'aux températures de l'ordre de T_{cr} l'énergie des électrons contenus dans la couche sphérique diffère de l'énergie de Fermi d'une quantité de l'ordre de kT_{cr}, de sorte que

$$kT_{cr} \approx \frac{p^2}{2m} - \frac{p_F^2}{2m} \approx \frac{(p+p_F)(p-p_F)}{m} \Delta p,$$

où $\Delta p = p - p_F$. Ainsi, l'indétermination sur la position des électrons $\Delta x \approx \hbar/\Delta p \approx mkT_{cr}/\hbar p_F$. L'ordre de grandeur de Δx est généralement de 10^{-4} cm (pour $T_{cr} \sim 10$ K). La valeur de Δx caractérise l'étendue de la région où sont localisés les électrons appartenant à une paire de Cooper, autrement dit elle caractérise l'« extension spatiale » de cette paire. On voit que ses dimensions sont près de 10 000 fois plus grandes que la distance moyenne entre les électrons de conduction dans un métal (qui est de l'ordre de 10^{-8} cm). Cela signifie qu'entre les électrons accouplés en paire de Cooper il se trouve un grand nombre d'autres électrons; c'est pour cela qu'on dit que les états des électrons appartenant à une paire de Cooper sont *faiblement correlés par rapport aux coordonnées* de l'espace ordinaire. La corrélation qui s'établit à des distances aussi grandes est inexplicable au point de vue classique et c'est un phénomène purement quantique. La *corrélation en impulsions est*, par contre, *forte*, puisque $p_1 = -p_2$.

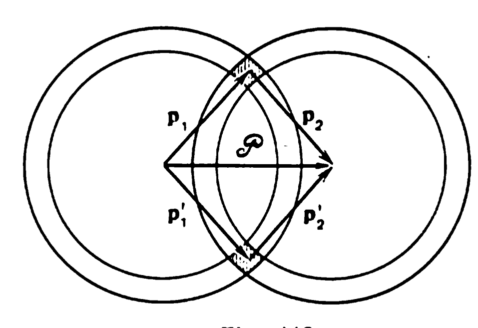


Fig. 118

5. L'existence d'un couplage des électrons dans les supraconducteurs se trouvant à une température $T < T_{cr}$ a été prouvée par des expériences directes relatives à la *quantification du flux magnétique.* Soit un anneau supraconducteur parcouru par un courant supraconducteur. Posons que les électrons se déplacent avec une vitesse v sur une trajectoire circulaire de rayon r. L'énergie du courant est donnée par l'expression $\mathcal{E} = \frac{1}{2c} I\Phi$, où I est l'intensité du courant et Φ le flux magnétique passant à travers le cercle considéré et créé par le courant. En notant N le nombre total d'électrons contenus dans l'anneau et T la période de révolution, $I = Ne/T = Nev/2\pi r$. Ainsi $\mathcal{E} = Nev\Phi/4\pi rc$. D'autre part, cette même énergie s'exprime

par $\mathscr{E} = {}^1/_2\, Nmv^2$. En identifiant ces deux expressions, on trouve $\Phi = 2\pi rcmv/e$. Si les électrons se déplacent sous forme de paires de Cooper, l'impulsion de chacune de ces paires est $p = 2mv$, de sorte que $\Phi = \pi rcp/e$. Or l'impulsion d'une paire de Cooper ne peut prendre que des valeurs quantifiées vérifiant la relation $pr = n\hbar = nh/2\pi$, où n est un nombre entier. Par conséquent,

$$\Phi = \frac{hc}{2e} n. \qquad (62.1)$$

Cette formule exprime la *quantification du flux magnétique* dans les supraconducteurs, le *quantum de flux magnétique* étant défini par l'expression

$$\Phi_0 = \frac{hc}{2e} = 2{,}07 \cdot 10^{-7}\,\text{Gs}\cdot\text{cm}^2. \qquad (62.2)$$

Une formule de cette forme avait été obtenue par F. London (1900-1954) dès 1950, i.e. avant l'élaboration de la théorie de supraconduction. Mais London trouva pour Φ_0 une valeur deux fois plus grande que celle donnée par la formule (62.2). Cela tient à ce qu'en 1950 on n'avait aucune notion du couplage d'électrons en paires, et de ce fait London utilisait pour l'impulsion la formule $p = mv$, au lieu de $p = 2mv$, qui a été utilisée pour établir la formule (62.2). Les expériences confirmèrent la justesse des formules (62.1) et (62.2) et par conséquent l'existence du phénomène de couplage des électrons. Dans une de ces expériences on a excité un courant supraconducteur dans un petit tube en étain de 1 cm de longueur et de diamètre intérieur $2R = 1{,}5 \cdot 10^{-3}$ cm. Vu la petitesse du diamètre, l'excitation d'un seul quantum de flux magnétique $\Phi_0 = \pi R^2 H$ conduit à l'apparition d'un champ magnétique d'intensité mésurable $H = 0{,}117$ Gs; le champ magnétique terrestre n'est que cinq fois environ plus grand. Ainsi, dans ces expériences la quantification du flux magnétique se manifeste à l'*échelle macroscopique*.

A ce propos il convient de noter les circonstances suivantes. On sait qu'on peut exciter un courant électrique non amorti dans un anneau supraconducteur. Une expérience de ce type a duré deux ans et demie et cependant on ne put déceler la moindre diminution de l'intensité du courant. A première vue, ce résultat peut sembler normal; puisque aucune chaleur de Joule n'est produite dans les supraconducteurs, il n'y a aucune raison pour que le courant diminue. En réalité la question n'est pas aussi simple. Dans un anneau supraconducteur les électrons se meuvent avec accélération et *devraient donc rayonner de l'énergie*, ce qui entraînerait un affaiblissement progressif du courant. Or des expériences prouvent qu'il n'y a aucune atténuation du courant. La contradiction est levée ici exactement de la même façon que fut levée la contradiction relative au rayonnement dans la théorie classique de l'atome. Pour éliminer la possibi-

lité de rayonnement Bohr introduisit le *postulat quantique* de l'existence d'états stationnaires de l'atome. De même un anneau supraconducteur parcouru par un courant ne rayonne pas par suite de la quantification du courant électrique ; cette quantification se manifeste cependant à une échelle macroscopique.

6. Les électrons appartenant à une paire de Cooper possèdent des impulsions de même module mais de sens opposé. En principe leurs spins peuvent être parallèles ou antiparallèles, puisque dans les deux cas le spin résultant de la paire de Cooper sera entier, et la paire de Cooper se comportera donc comme un boson et non comme un fermion. Pour un système à spins parallèles le spin de la paire de Cooper sera égal à zéro et le système sera instable, car c'est le système à spins antiparallèles qui est stable et qui sera donc réalisé. *Les paires de Cooper se comportent donc comme des bosons*, et le principe d'exclusion de Pauli ne s'y applique donc pas, ce qui lève la difficulté signalée au début de ce paragraphe. Les paires de Cooper constituent un « fluide » pouvant effectuer un mouvement superfluide entre les ions du réseau cristallin. C'est ça le phénomène de supraconduction. Examinons cette question de plus près.

A $T = 0$ toutes les paires de Cooper se trouvent à l'état fondamental et leurs impulsions sont nulles. Pour $T < T_{cr}$ la majorité des paires de Cooper se trouvent à l'état fondamental et constituent une *collectivité* de bosons et se trouvent donc toutes *dans le même état quantique*. Si on provoque la rupture d'une paire de Cooper et si on élimine de la collectivité des paires l'électron libéré, on verra apparaître un état excité ayant une énergie plus grande que celle de l'état initial. Ce qui importe c'est que ces deux états sont séparés par une *fente énergétique* de largeur finie, qui ne comporte aucun état quantique permis pour le système. Quoique la largeur de la fente soit très petite (à $T = 0$ elle est de l'ordre de 10^{-3} à 10^{-4} eV), son existence entraîne un changement qualitatif dans le comportement de tout le système électronique. Bien que certaines paires de Cooper soient désunies par suite de leurs interactions avec les vibrations du réseau, tant que $T < T_{cr}$, les électrons provenant de ces paires rompues ne peuvent traverser la fente énergétique, car ils ne possèdent pas une énergie suffisante pour cela. Un électron devenu « célibataire » se cherche un nouveau partenaire parmi les autres électrons libérés pour constituer une nouvelle paire de Cooper d'impulsion et de spin nuls. Or, vu l'indiscernabilité des électrons, ces processus d'échange ne peuvent conduire à un nouvel état et à la destruction de la collectivité. Aux températures $T < T_{cr}$ la collectivité se comporte et se déplace comme un tout. Mais lorsque $T > T_{cr}$ les électrons provenant des paires rompues peuvent surmonter la fente énergétique et s'échappent de leur collectivité qui s'en trouve décimée.

C'est l'existence d'une fente de largeur finie dans le spectre énergétique de l'état supraconducteur qui explique la disparition de tou-

te résistance au passage d'un courant électrique. En l'absence de courant électrique les impulsions de toutes les paires de Cooper sont nulles, puisque l'impulsion d'un électron étant égale à $+\boldsymbol{p}$, celle de l'autre est égale à $-\boldsymbol{p}$. Lorsqu'on applique un champ électrique $\boldsymbol{E}$ les deux électrons d'une paire acquièrent la même impulsion supplémentaire $\boldsymbol{p}'$, et leurs impulsions totales deviennent égales respectivement à $\boldsymbol{p} + \boldsymbol{p}'$ et $-\boldsymbol{p} + \boldsymbol{p}'$; l'impulsion de la paire devient donc égale à $2\boldsymbol{p}'$. La collectivité de paires de Cooper se met en mouvement d'un seul tenant avec une vitesse déterminée correspondant à cette valeur de l'impulsion, ce qui marque l'apparition d'un courant électrique. Tant que l'intensité de courant n'est pas grande, l'existence de la fente énergétique protège la collectivité des paires de Cooper contre la destruction et le courant ne rencontre aucune résistance. La collectivité des paires de Cooper se trouve alors dans l'état de passage du courant électrique, et, à l'instant où s'établit le courant continu, le champ électrique $\boldsymbol{E}$ devient nul dans le supraconducteur. Bien entendu, la collectivité de paires de Cooper en mouvement continue à être le siège d'interactions électrons-phonons qui assurent la stabilité de l'état correspondant au transport du courant. Lorsque l'intensité du courant devient assez grande, la collectivité s'en trouve dispersée et la supraconduction cesse d'exister.

A la température du zéro absolu tous les électrons du supraconducteur sont liés dans des paires de Cooper. Lorsque la température s'élève, le nombre de paires diminue et ce pour deux raisons. Premièrement, à mesure que la température, s'élève, augmente la probabilité pour qu'un électron se détache d'une paire de Cooper et franchisse la fente. Deuxièmement, le nombre de processus conduisant à la formation de paires diminue parce que diminue la fréquence des échanges de phonons entre les électrons déterminant leur attraction mutuelle. En conséquence, la largeur de la fente diminue, ce qui accélère la décomposition des paires de Cooper. A $T = T_{cr}$ la fente disparaît, le nombre de paires de Cooper tombe à zéro et la collectivité de ces paires se trouve anéantie. En même temps la supraconduction disparaît et, à $T > T_{cr}$ le métal redevient un conducteur normal.

7. La similitude entre les phénomènes de supraconduction et de superfluidité se manifeste dans la prédiction de la superfluidité de $^{3}_{2}$He. Comme les atomes de cet isotope ont un spin demi-entier, on supposa d'abord qu'il ne pouvait devenir superfluide. Aussitôt après l'élaboration de la théorie de supraconduction on envisagea la possibilité de formation de paires d'atomes $^{3}_{2}$He analogues aux paires de Cooper. Comme les paires d'atomes $^{3}_{2}$He possèdent un spin entier, elles se comportent en bosons, ce qui permet d'envisager la possibilité de la superfluidité de $^{3}_{2}$He, qui fut bientôt découverte expérimentalement (cf. § 61).

8. La théorie fournit aussi une explication de l'effet Meissner-Ochsenfeld (cf. t. III, § 80). On sait que cet effet consiste en ce que

le champ magnétique ne pénètre pas dans un supraconducteur massif ou s'en trouve expulsé. Pour expliquer simplement cet effet on considérera un supraconducteur en forme de cylindre suffisamment long pour pouvoir exclure les effets de bord. Lorsque le champ magnétique est parallèle à l'axe du cylindre, il suffit de remarquer que le courant supraconducteur ne rencontre aucune résistance électrique, même s'il circule uniquement en surface (plus exactement dans une mince couche périphérique). Il faut noter encore que de tous les états envisageables d'un supraconducteur ne sera réalisé que celui qui correspond à la plus petite énergie.

Posons $T > T_{cr}$ et plaçons le supraconducteur dans un champ magnétique H permanent et homogène. On peut parfaitement admettre que le champ pénètre dans le supraconducteur en s'y répartissant uniformément de façon à rester homogène. En vertu des conditions aux limites, les intensités du champ magnétique H à l'intérieur et l'extérieur du cylindre supraconducteur doivent être égales. L'énergie magnétique du système sera donc égale à

$$\mathscr{E}_1 = \mathscr{E}_{ex} + \mathscr{E}_i,$$

où $\mathscr{E}_{ex}$ est l'énergie magnétique localisée à l'extérieur du cylindre et $\mathscr{E}_i \doteq \frac{1}{8\pi}\mu V H^2$ l'énergie magnétique localisée à l'intérieur du cylindre (V — volume du cylindre, μ — perméabilité magnétique du supraconducteur). Supposons maintenant que des courants circulaires circulent à la surface du cylindre et créent un champ magnétique H_1 uniforme dans son volume. Comme ces courants ne créent aucun champ à l'extérieur du cylindre, on peut représenter l'énergie magnétique par l'expression

$$\mathscr{E}_2 = \mathscr{E}_{ex} + \frac{\mu}{8\pi} V (H + H_1)^2,$$

l'énergie $\mathscr{E}_{ex}$ localisée dans l'espace extérieur sera la même que dans le cas précédent. L'énergie $\mathscr{E}_2$ devient minimale $\mathscr{E}_{2\min} = \mathscr{E}_{ex}$ lorsque $H + H_1 = 0$. C'est cet état d'énergie minimale qui doit se réaliser à la place de $\mathscr{E}_1$, puisque rien n'empêche le courant de se redistribuer dans le volume du supraconducteur. Il s'ensuit que dès qu'un cylindre supraconducteur est introduit dans un champ magnétique uniforme, on voit apparaître des courants circulaires superficiels supraconducteurs qui détruisent le champ magnétique à l'intérieur du cylindre. C'est l'effet Meissner-Ochsenfeld qui de toute évidence ne peut se manifester que si $T < T_{cr}$ puisqu'au-dessus de T_{cr} aucun courant superficiel supraconducteur ne peut circuler.

9. Dans ce qui précède il n'a pas été tenu compte de l'*énergie superficielle* existant à la surface de séparation de la phase supraconductrice avec la phase normale. Cette énergie dépend de la profondeur

finie de pénétration dans la phase supraconductrice du champ magnétique régnant dans la phase normale, de l'attraction mutuelle des électrons d'une paire de Cooper, de l'existence d'une fente énergétique entre les phases normale et supraconductrice, etc. Cette énergie superficielle peut être positive ou négative et c'est ce qui suggéra à A. A. Abrikossov (né en 1928) à diviser tous les supraconducteurs en *supraconducteurs de première* et *de deuxième espèce.* L'énergie superficielle est positive dans les supraconducteurs de première espèce et négative dans les supraconducteurs de deuxième espèce. Parmi les supraconducteurs de première espèce se classent la majorité des métaux purs, tandis que la grosse majorité des alliages et nombre de métaux purs dopés par des impuretés sont des supraconducteurs de deuxième espèce. L'effet Meissner-Ochsenfeld s'observe dans tous les supraconducteurs de première espèce mais pas toujours dans les supraconducteurs de deuxième espèce. Ces derniers peuvent se trouver dans un *état supraconducteur* ou dans un *état mixte*: lorsqu'ils se trouvent à l'état supraconducteur ils manifestent l'effet Meissner-Ochsenfeld et lorsqu'ils se trouvent à l'état mixte cet effet est absent.

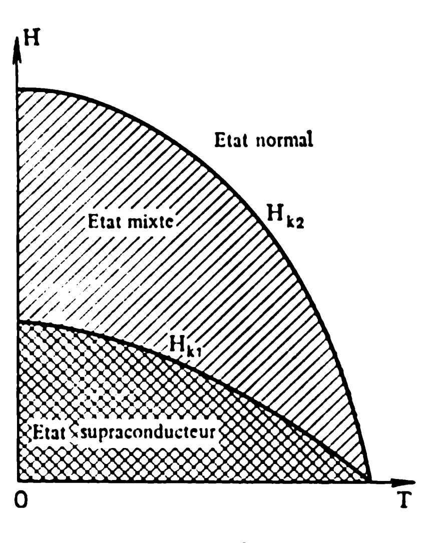


Fig. 119

Sur la figure 119 la courbe $H = H_{cr_1}(T)$ permet de déterminer l'intensité du champ critique à laquelle les phases supraconductrice et mixte sont en équilibre; de même la courbe $H = H_{cr2}(T)$ correspond à l'équilibre entre la phase supraconductrice et la phase normale. Le domaine des températures et des intensités du champ magnétique auxquelles le métal se trouve à l'état supraconducteur est indiqué par des hachures croisées, le domaine de l'état mixte est dénoté par des hachures simples et le domaine de l'état normal n'est pas hachuré. L'état mixte n'existe pas dans les supraconducteurs de première espèce.

Comme dans tout supraconducteur doit s'instaurer l'état auquel correspond le *minimum* de l'énergie totale, l'énergie superficielle y comprise, on conçoit l'apparition de l'état mixte. Lorsqu'un supraconducteur se trouve à l'état mixte, le champ magnétique extérieur y pénètre sous forme des *domaines filiformes de section droite finie.* La section droite est nécessairement finie puisque un domaine étant occupé par le champ magnétique ce dernier doit pénétrer à une cer-

taine profondeur dans l'espace environnant occupé par la phase supraconductrice ; ce processus se caractérise par une profondeur de pénétration finie. Le corps est transpercé par des domaines filiformes traversés par des flux magnétiques, ces domaines étant séparés les uns des autres par des régions supraconductrices, à condition que l'intervalle entre les domaines voisins est supérieur au double de la profondeur de pénétration du champ magnétique dans le supraconducteur.

Il est essentiel que le flux magnétique passant à travers la section droite des domaines normaux est *quantifié*. Il est énergétiquement avantageux que chaque domaine ne soit traversé que par un *seul quantum* de flux magnétique. Pour le démontrer considérons deux domaines de rayon r, traversés chacun par un quantum de flux magnétique. Le flux magnétique total traversant les deux domaines est égal à $2\pi r^2 H$. Supposons maintenant que les deux domaines soient réunis en un seul de rayon R. Le même flux sera représenté alors par $\pi R^2 H$. En comparant les deux expressions, on trouve $R = r\sqrt{2}$. Par conséquent, la longueur du cercle de la section droite du domaine de rayon R est $2\pi R = 2\pi r\sqrt{2}$, tandis que la somme des longueurs des cercles des sections droites des deux domaines initiaux est égale à $2\pi r \cdot 2$. Ce résultat montre que la réunion de deux domaines en un seul *diminue la surface latérale* séparant la phase normale de la phase supraconductrice. Ceci conduit à l'accroissement énergétiquement désavantageux de l'énergie superficielle, car elle est négative. On voit que quoique le corps soit traversé par le champ magnétique il reste supraconducteur grâce à l'existence de régions supraconductrices disposées entre les domaines normaux. Lorsque l'intensité du champ magnétique augmente le nombre de domaines normaux augmente et le volume des régions supraconductrices diminue. Finalement le champ magnétique occupe tout le volume et la supraconduction disparaît.

Les alliages supraconducteurs ayant de grandes valeurs du champ magnétique critique H_{cr2} trouvent de nombreuses applications pour la fabrication de solénoïdes produisant des champs magnétiques de très grande intensité (100 000 Gs et au-dessus).

INDEX DES NOMS

INDEX DES MATIÈRES

TABLE DES MATIÈRES

Chapitre IV. L'ÉQUATION DE SCHRÖDINGER ET LA QUANTIFICATION

Chapitre V. NOUVELLES ÉTAPES DE L'ÉLABORATION DE LA MÉCANIQUE QUANTIQUE ET SPECTRES

Chapitre VI. SYSTÈMES ATOMIQUES À PLUSIEURS ÉLECTRONS

Chapitre VII. EXEMPLES D'EFFETS QUANTIQUES MACROSCOPIQUES

À NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur le contenu de ce livre, sa traduction et sa présentation, ainsi que toute autre suggestion.

Notre adresse:
2, Pervi Rijski péréoulok,
Moscou, I-110, GSP, U.R.S.S.

Imprimé en Union Soviétique